AD_____

Award Number: DAMD17-03-1-0270

TITLE: Genetic and Hormonal Risk Factors for Prostate Cancer in African American Men

PRINCIPAL INVESTIGATOR: Aruna Sarma, Ph.D.

CONTRACTING ORGANIZATION: The University of Michigan
Ann Arbor, MI 48109

REPORT DATE: May 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-05-2006	Annual	1 MAY 2005 - 30 APR 2006

4. TITLE AND SUBTITLE
Genetic and Hormonal Risk Factors for Prostate Cancer in African American Men

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-03-1-0270

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Aruna Sarma, Ph.D.

E-mail: asarma@umich.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The University of Michigan
Ann Arbor, MI 48109

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
Original contains color plates: All DTIC reproductions will be in black and white.

14. ABSTRACT

African American men are at greater risk for developing and dying from prostate cancer compared to white men. This disparity is likely due to a number of factors including environmental and genetic factors. The Flint Men's Health Study (FMHS) was established in 1995 as a population-based case-control study of African American men aged 40-79 residing in Genesee County, Michigan. The initial sample consisted of 730 men who completed an in-home interview consisting of potential risk factors for prostate cancer; medical history; and demographic data. 431 men provided a blood sample and 369 men who were determined to be free of cancer completed a comprehensive urologic exam. Additionally, 119 cases of prostate cancer have been identified from the same population. Studies have suggested a role for hormones and genetics in cancer incidence. However, studies have been completed in white populations and results have been conflicting. The objective of this study is to more clearly delineate the potential role(s) of selected hormones and growth factors in prostate cancer development.

15. SUBJECT TERMS
Prostate Cancer, Association Studies, Polymorphisms, Androgens

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					USAMRMC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	31	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

INTRODUCTION

Prostate cancer is the most common cancer and the second leading cause of cancer deaths in U.S. men. African American men are at greater risk for both developing and dying from prostate cancer compared to white men. Research based on racial differences in prostate cancer has resulted in important discoveries that have just begun to unravel the complex biologic mechanisms for prostate cancer. The paucity of data on prostate cancer from population-based samples of African American men highlights a gap in our current understanding of the disease. The goal of this proposed study is to determine whether or not there are differences in 1) various circulating hormone and growth factor levels and 2) the prevalence of genetic polymorphisms between African American men with prostate cancer and those free of disease. We will also measure the associations between genetic polymorphisms and circulating hormone levels and their subsequent impact on prostate cancer risk in African American men. This information will enable us to better understand the role of hormonal and genetic risk factors in the disease process. We will make use of the unique opportunity presented by the availability of an ongoing epidemiologic study of community-dwelling African American men: the Flint Men's Health Study. We will accomplish this through the following Specific Aims:

Specific Aim 1: To evaluate the associations between *circulating hormone* and *growth factor* levels and *prostate cancer* using samples from a population-based cohort of African American men.

Specific Aim 2: To evaluate the associations between *genetic polymorphisms* and *prostate cancer* using samples from a population-based cohort of African American men.

a. To evaluate differences in the prevalence of common *genetic polymorphisms* between men diagnosed with prostate cancer and disease free controls.

b. To evaluate differences in the prevalence of common *genetic polymorphisms* between men diagnosed with various stage and grade prostate cancer.

Specific Aim 3: To measure the associations between genetic polymorphisms and circulating hormone levels and their subsequent impact on prostate cancer risk in African American men.

The completion of these aims will lead to new insights into differences and/or similarities in the prevalence of hormonal and genetic correlates of prostate cancer between cases and controls. These insights will provide the direction for the next sets of studies to better define the etiology of prostate cancer in African American men, its natural history and clinic course, as well as potential targets for intervention.

BODY

The following include research accomplishments associated with tasks outlined in our approved Statement of Work.

Task 1. **To perform laboratory assays for the following hormones:**

Months 1-12
Total Testosterone
Free Testosterone
Androstenedione
Dihydroepiandrosterone sulfate
Serum hormone binding globulin
Androstenediol glucuronide
Sodium

Months 12-24
Estradiol
Estrone Sulfate
Insulin-like Growth Factor-1
Insulin-like Growth Factor Binding Protein-3

This task has been completed and included sending aliquots of over 450 serums to the appropriate laboratories, entering the data into spreadsheets, and data cleansing. Specifically, circulating hormone levels listed above were quantified by performing laboratory assays at the University of Michigan Medical Center Reproductive Sciences and Clinical Chemistry Labs. Total IGF-1, IGFBP-3 and AAG were measured by a commercially available enzyme-linked immunosorbent assays (ELISA) (Diagnostic Systems Laboratory, Webster, Texas) in Dr. Jaffe's laboratory. Inter-assay and intra-assay coefficients of variation were as follows: IGF-1: 4%,6%; IGFBP-3: 6%,9%; and AG: 5% and 11%. All other hormones were measured using commercially available chemiluminescent immunoassays (Bayer Diagnostics, Pittsburgh, PA) Inter-assay and intra-assay coefficients of variation, respectively, were as follows: DHEAS:18.14%, 12.53%; TT:8.68%, 6.82%; FT:6.5%, 7.3%; SHBG:18.95%, 10.31%; Androstenedione:11.6%, 6 %; E2:10.21%, 6.375%; and Estrone:10.27%, 10.93%.

Task 2. **To perform DNA analyses to examine the following genotypes:**

Months 1-12	Months 12-24
LHB	CYP19
HSD3B2	CYP3A4
CYP17	IGF1
HSD17B2	

We had initially planned to perform our SNP assays using either ABI PRISM® 7700 Sequence Detection System which uses TaqMan® assays or an ABI 3100 PRISM®Genetic Analyzer which employs a 16 capillary electrophoresis system. However the University of Michigan Comprehensive Cancer Center cDNA Core recently purchased the ABI PRISM® 7900 Sequence Detection System which has the capacity to perform high-throughput SNP detection using 384-well plates. We have performed a number of pilot experiments and have learned that SNP assays using the ABI 7900 are very sensitive to DNA concentration. We have optimized our assays and the failure rate is now less than 5%. This new genotyping platform has the reduced the estimated cost per genotype from $2.00 to less than $0.50. This allowed us to perform additional genotyping. Additionally in the face of new analytical methods for SNP selection and analyses we have revised our genotype list from the one listed above.

In summary, we completed in our lab genotyping for the PSA and Insulin genes using methods described above. Genotyping for the AR, CYP11A1, SRD5A2 genes were completed at the Mayo Clinic while assays for the MSR1 genotypes were completed at Wakeforest University. A manuscript describing data from the MSR1 analyses was published in *Cancer Research(1)* (See Appendix) and PSA, Insulin and chromosome 8q24 genotyping findings have been published in *Prostate Cancer and Prostatic Diseases,* the *Prostate,* and *Nature Genetics* journals, respectively(2-4) (See Appendix). Data quality assessment, cleansing and analyses of AR, CYP11A1 and SRD5A2 are currently in progress. We decided to outsource our genotyping for the remainder of the SNPs in the interest of time and money. The following SNPs were selected and assays completed. Data are now being analyzed.

IGFBP-3	rs3110697	**CY3A4**	rs12333983
	rs3793345		rs2242480
CYP17	rs10883783		rs2740574
	rs6163	**CYP19**	rs7000519
	rs6162	**IGF-1**	CA repeat
	rs743572		

Task 3. **Interim statistical analyses of data obtained from hormone assays and DNA will be performed periodically as assays are performed. (Months 1-24)**
For years 1-2, we completed preliminary statistical analyses of proposed serum hormone data and statistical analyses of genotyping have been completed for the following genes: MSR1, PSA, Insulin and Allele -8 of microsatellite DG8S737. Analyses of all data regarding the other SNPs listed above are currently in progress. Additionally, more comprehensive analyses and manuscript preparation for various serum hormone data among cases and additional analyses assays for total and free testosterone are underway.

Task 4. **Final analyses and report writing (Months 24-36)**
a. Final analyses of data from hormone assays and genetic polymorphisms will be performed.
b. Final manuscripts will be prepared and submitted.

Four manuscripts have been published. Final analyses of hormone data are currently in progress and analyses and report writing of remaining genetic polymorphisms will be completed in the final year.

KEY RESEARCH ACCOMPLISHMENTS

- A significant difference in median values of sex hormones was observed between serum from African American prostate cancer cases and disease-free controls
- Selected rare macrophage scavenger receptor 1 gene mutations were observed to be associated with increase prostate cancer susceptibility among African American men
- Polymorphisms in the prostate specific antigen gene promoter did not predict serum prostate specific antigen levels nor prostate cancer risk in African American men
- Insulin gene *PstI* CC genotype was positively associated with prostate cancer diagnosis and lower grade and stage of prostate cancer in African-American men. INS *PstI* genotype was not associated with later age of diagnosis.
- Allele -8 of the microsatellite DG8S737 was associated with prostate cancer in both European and African-American men. A greater population attributable risk in African-American men suggests that this variant may contribute to higher incidence of prostate cancer in African-American men than in men of European ancestry.
- Preliminary data suggest potential associations between the selected SNPs and prostate cancer diagnosis

REPORTABLE OUTCOMES

* Kruskal-Wallis Test.

In this first 3 years of funding, we have made significant progress towards completing laboratory studies required to carry out analyses described in Specific Aims 1, 2 and 3. Table 1 reports initial comparisons of mean serum hormone concentrations by prostate cancer status. Control

Table1 . Median (25%, 75%) values of sex hormone-related factors according to prostate cancer status.

Hormone	Prostate Cancer (n=124)	Disease-free controls (n=406)	P-value*
Androstenedione (ng/ml)	6.28 (3.90, 9.06)	1.00 (0.80, 1.30)	<0.0001
Estradiol (pg/ml)	36.70 (26.75, 43.5)	28.90 (22.90, 36.40)	<0.0001
Estrone (ng/ml)	1.44 (1.04, 1.87)	1.91 (1.25, 2.94)	<0.0001
SHBG (nM)	42.15 (30.20, 56.60)	29.70 (21.50, 43.10)	<0.0001
Total testosterone (ng/dl)	445.50 (321.46, 611.62)	559.77 (413.88, 757.10)	<0.0001
IGF-1 (ng/ml)	233.66 (183.55, 288.16)	59.00 (41.00, 83.00)	<0.0001
IGFBP-1(ng/ml)	39.62 (31.89, 46.52)	3.32 (2.60, 4.33)	<0.0001

serums were evaluated for quality control purposes over the last year and temporal and batch variations were found on the original samples calling into question the validity of the original control hormone values. A separate manuscript describing distributions of hormone concentrations among prostate cancer cases only is in preparation. Among a pristine sample of serum in controls (deemed of high quality and void of temporal or batch variations), total and free testosterone concentrations will be re-assayed. Once these evaluations are complete, additional analyses will be performed to examine potential confounders in the relationship between these total and free testosterone concentrations and the risk of prostate cancer and to determine whether significant differences in serum hormone concentrations exist by stage and grade of disease.

In addition, several manuscripts have been completed (4 published) examining the association between various SNPs and risk of prostate cancer in African-American men. Below are the published abstracts for the three completed analyses. Preliminary data on remainder of SNPs are depicted in Table 2 below.

"Germ-line Mutations of the Macrophage Scavenger Receptor 1 Gene: Association with Prostate Cancer Risk in African American Men." *Cancer Research. 63, 3486-89. 2003.*

Both rare germ-line mutations and common sequence variants of the macrophage scavenger receptor 1 (*MSR1*) gene have recently been implicated as potential prostate cancer susceptibility factors. However, existing studies are limited by the referral-based nature of samples and a paucity of African-American participants. In this context, we evaluated the association of germ-line mutations and common *MSR1* sequence variants with prostate cancer risk in a case control study of a community-based sample of 134 African-American men with prostate cancer and 340 unaffected controls. In our sample, the rare Asp174Tyr missense change was identified nearly twice as frequently in men with prostate cancer (6.8%) compared with unaffected controls (3.6%; $P = 0.14$). Moreover, significantly different allele frequencies between cases and controls were observed for one of the sequence variants, IVS5–59 ($P = 0.02$). Taken together, our results provide some additional support for the hypothesis that selected, rare *MSR1* mutations are associated with increased prostate cancer susceptibility among African-American men. (See Appendix)

"Polymorphisms in the Prostate Specific Antigen Gene Promoter Do Not Predict Serum Prostate Specific Antigen Levels in African American Men." *Prostate cancer and prostatic diseases 9, 50 -55, 2006.*

A major problem with the use of serum PSA in predicting prostate cancer risk is the considerable intra-individual variability of such measurements. Cramer et al. (2003) identified a set of single nucleotide polymorphisms (SNPs) in the upstream regulatory region of the PSA gene that were each associated with increased promoter activity and serum PSA, further suggesting that genotyping these SNPs could be useful in improving the predictive value of PSA screening. In order to replicate this finding, DNA samples from 475 African American men were genotyped for the same SNPs and no association was observed with either serum PSA level or prostate cancer diagnosis. (See Appendix)

"INS PstI Polymorphism and Prostate Cancer Risk in African-American Men: The Flint Men's Health Study" *Prostate 65, 83 -87, 2005.*

BACKGROUND: Both prostate cancer and diabetes mellitus are common disease in African-American men. High insulin levels and insulin resistance have been implicated in prostate cancer development, which has prompted a recent investigation of a possible role for germline variation in the insulin gene (INS) and prostate cancer risk.

METHODS: Four hundred sixty-six African-American men with and without prostate cancer from the Flint Men's Health Study were typed for the INS Pst1 genotype using restriction digest and direct sequencing. An association between the Pst1 genotype and prostate cancer was examined using crude and age-adjusted logistic regression models.
RESULTS: African-American men who were homozygous for the INS Pst1 CC genotype were 1.59 times more likely to be diagnosed with prostate cancer compared to men with the TT or TC genotypes (95%CI=0.93, 2,72). The association appeared stronger among diabetics compared to non-diabetics; however this observation was not statistically significant.
CONCLUSIONS: Our study, take together with the report of [Ho et al. Br J Cancer 88:263-269, 2003], suggests that the INS Pst1 CC genotype is associated with prostate cancer risk in African-American men. Germline variation in the INS gene should be more fully explored in multiethnic studies to elucidate the molecular variant(s) associated with prostate carcinogenesis. (See Appendix)

"A Common Variant Associated with Prostate Cancer in European and African Populations" *Nature Genetics, Advanced Online Publication, 1-7, 2006.*

With the increasing risk of prostate cancer, identifying common genetic variants that confer risk of disease is important. In this report, samples of DNA from the FMHS were used in a report on a variant on chromosome 8q24, a region initially identified through a study of Icelandic families. Allele -8 of the microsatellite DG8S737 was associated with prostate cancer in three case control series of European ancestry from Iceland, Sweden and the US. The estimated odds ratio (OR) of the allele is 1.62 ($P=2.7 \times 10^{-11}$). About 19% of affected men and 13% of the general population carry at least one copy, yielding a population attributable risk (PAR) of ~8%. The association was also replicated in an African-American case-control group (FMHS) with a similar OR, in which 41% of affected individuals and 30% of the population are carriers. This leads to a greater estimated PAR (16%) that may contribute to higher incidence of prostate cancer in African American men of European ancestry. (See Appendix)

Table 2. Genotype Information for Selected SNPS: Preliminary Data

SNPs	Gene	Chromosome	Position	Nucleotide Change	Frequency of Variant Allele in Cases	Frequency of Variant Allele in Controls
rs3110697	*IGFBP3*	7p13-p12	Intron 3	C to T	59.4	63.3
rs3793345	*IGFBP3*	7p13-p12		A to G (T to C NCBI)	32.3	40.4
rs10883783	*CYP17*	10q24.3	Intron 7	A to T (T to A NCBI)	30.8	31.4
rs6163	*CYP17*	10q24.3	Exon 1 (Codon 65)	G to T	57.9	66.4
rs6162	*CYP17*	10q24.3	Exon 1 (Codon 46)	C to T	58.7	66.9
rs743572	*CYP17*	10q24.3	5'-UTR	T to C	57.8	66.9
rs12333983	*CYP3A4*	7q21.1	3' UTR	A to T (T to A NCBI)	89.7	85.5
rs2242480	*CYP3A4*	7q21.1	Intron 10	A to G (T to C NCBI)	65.4	54.8
rs2740574	*CYP3A4*	7q21.1	5' UTR (-293)	G to A	49.6	46.8
rs700519	*CYP19*	15q21.1	Exon 6 (codon 264)	C to T	30.1	30.8
rs9282858	*SDR5A2*	2p23	Exon 1 (codon 49)	G to A	0.8	1.4
rs523349	*SDR5A2*	2p23	Exon 1 (codon 89)	C to G	93.2	90.9
rs9332964	*SDR5A2*	2p23	Exon 4 (codon 227)	G to A	100	100

CONCLUSIONS

Given the overall increased incidence of and mortality due to prostate cancer in African-American men, these men represent an appropriate population for the study of the associations between circulating hormones and genetic polymorphisms and prostate cancer risk. Our goal is to identify hormonal factors and genetic markers that can be used to stratify African-American men who are at risk for developing disease as well as those who progress to more severe disease. Over the next 12 months, we will complete our analyses of the hormone and genotype data with respect to prostate cancer risk.

REFERENCES (Bibliography of completed manuscripts)

(1) Miller D, Zheng SL, Dunn RL, Sarma AV, Montie JE, Lange EM et al. Germ-line Mutations of the Macrophage Scavenger Receptor 1 Gene: Association with Prostate Cancer Risk in African-American Men. Cancer Research 2003; 63:3486-3489.

(2) Claeys GB, Sarma AV, Dunn RL, Zuhlke KA, Beebe-Dimmer J, Montie JE et al. INS Pst1 Polymorphism and Prostate Cancer in African-American Men. The Prostate 2005; 65:83-87.

(3) Beebe-Dimmer J, Lange LA, Cain JE, Lewis RC, Ray AM, Sarma AV et al. Polymorphisms in the Prostate-Specfic Antigen Gene Promoter Do Not Predict Serum Prostate-Specific Antigen Levels in African-American Men. Prostate Cancer and Prostatic Diseases 2006; 9:50-55.

(4) Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarrsson B et al. A Common Variant Associated with Prostate Cancer in European and African Populations. Nature Genetics 2006; Advanced Online Publication:1-7.

[CANCER RESEARCH 63, 3486–3489, July 1, 2003]

Advances in Brief

Germ-line Mutations of the Macrophage Scavenger Receptor 1 Gene: Association with Prostate Cancer Risk in African-American Men[1]

David C. Miller, S. Lilly Zheng, Rodney L. Dunn, Aruna V. Sarma, James E. Montie, Ethan M. Lange, Deborah A. Meyers, Jianfeng Xu, and Kathleen A. Cooney[2]

Departments of Urology [D. C. M., R. L. D., A. V. S., J. E. M., K. A. C.] and Internal Medicine [K. A. C.], University of Michigan, Medical School, Ann Arbor, Michigan 48109-0946, and Department of Public Health Sciences and Center for Human Genetics, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157-1063 [S. L. Z., E. M. L., D. A. M., J. X.]

Abstract

Both rare germ-line mutations and common sequence variants of the macrophage scavenger receptor 1 (*MSR1*) gene have recently been implicated as potential prostate cancer susceptibility factors. However, existing studies are limited by the referral-based nature of samples and a paucity of African-American participants. In this context, we evaluated the association of germ-line mutations and common *MSR1* sequence variants with prostate cancer risk in a case control study of a community-based sample of 134 African-American men with prostate cancer and 340 unaffected controls. In our sample, the rare Asp174Tyr missense change was identified nearly twice as frequently in men with prostate cancer (6.8%) compared with unaffected controls (3.6%; $P = 0.14$). Moreover, significantly different allele frequencies between cases and controls were observed for one of the sequence variants, IVS5–59 ($P = 0.02$). Taken together, our results provide some additional support for the hypothesis that selected, rare *MSR1* mutations are associated with increased prostate cancer susceptibility among African-American men.

Introduction

There is a growing body of molecular and genetic epidemiological evidence that implicates the short arm of chromosome 8 (8p22–23) as the location of one or more genes important in the development of adenocarcinoma of the prostate (1, 2). Most recently, the *MSR1*[3] gene has been proposed as an etiologic link between germ-line alterations in 8p and prostate carcinogenesis (3, 4). Xu *et al.* identified several rare germ-line mutations of the *MSR1* gene that cosegregated with prostate cancer among families affected with HPC. Moreover, at least one of the germ-line mutations was associated with an increased risk of sporadic prostate cancer among African-American men (4). In a subsequent study of men of European descent, the same authors examined five common sequence variants of *MSR1* and reported significantly different allele frequencies for each of the five variants among men diagnosed with prostate cancer compared with unaffected controls. Notably, the association of the common sequence variants with prostate cancer risk was independent of the presence of rare germ-line mutations (3).

The composite results of these studies provide provocative data in support of *MSR1* as a prostate cancer susceptibility gene. However, the generalizability of these findings is limited by a lack of African-Americans participants. Given that African-American men have both a higher incidence and mortality from prostate cancer compared with Caucasian men in the United States, characterization of genetic risk factors in this patient population is an important public health initiative, and further study of a potential role for *MSR1* is warranted (5). The aim of this study is to further evaluate the association between genetic variation in the *MSR1* gene and prostate cancer susceptibility among African-American men.

Materials and Methods

Subjects. Both cases and controls were recruited as part of the FMHS. Informed consent was obtained from each study participant, and all research protocols were approved by the Institutional Review Board at the University of Michigan Medical School. As described previously, disease-free controls, aged 40–79, were identified from a probability sample of African-American men in the city of Flint, Michigan or in neighboring Beecher Township (Genesee County; Ref. 6). A complete urological history and physical examination, including PSA testing, was performed to exclude the diagnosis of prostate cancer. Participating community urologists used the PSA values in conjunction with other clinical data to determine the need for biopsy; in general, a PSA value of >4 ng/ml indicated the need for biopsy. DNA was available for genetic sequencing for 345 unaffected men; however, the DNA was insufficient for 5 individuals. Thus, our final control sample consists of 340 disease-free African-American males.

Prostate cancer case recruitment from the same community was initiated in 1999 and completed in July 2002. Participation of cases required: (*a*) an epidemiological interview; (*b*) a review of the hospital and registry records for information on tumor stage, Gleason Score, prediagnosis PSA, and type of therapy; and (*c*) provision of a blood sample for DNA and freezer storage of serum and plasma. After excluding two cases with insufficient DNA, our final case sample included 134 African-American men, aged 40–79, that had been diagnosed with prostate cancer between 1995 and 2002. For both cases and controls, genomic DNA was isolated from whole blood by the use of the Puregene kit (Gentra Systems, Inc., Plymouth, MN).

Sequence Analysis. Five common sequence variants and five recently reported rare germ-line mutations were analyzed for 134 cases and 340 unaffected controls. The five rare mutations were identified during screening for sequence variants of *MSR1* in germ-line DNA samples from individuals with HPC (4). Four are missense mutations (Ser41Tyr, Asp174Tyr, Gly294Glu, and Pro36Ala), and one is a nonsense change (Arg293X). The five common sequence variants genotyped have been described previously and include an SNP in the promoter sequence (PRO3), a 15-bp insertion/deletion variant in intron 1 (INDEL1), an SNP located in intron 5 (IVS5-59), a missense mutation in exon 6 (P275A), and a 3-bp insertion/deletion in intron 7 (INDEL7; Ref. 3). The method of identification and positions of the five sequence variants have been reported elsewhere (3).

Statistical Analysis. Bivariate comparisons of mutation and allele frequencies among cases and controls were carried with χ^2 analysis or Fisher's exact test. Logistic regression models were used to test the association between common variants and disease status. These models were age adjusted to account for the possibility that some of the controls may later become diagnosed as cases. To avoid bias, age was calculated based on the same date for all cases and controls. This date was the most recent follow-up date from the

Received 3/7/03; accepted 5/6/03.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by USPHS Grant P50 CA69568 and the Department of Urology, University of Michigan Medical School.

[2] To whom requests for reprints should be addressed, at Departments of Internal Medicine and Urology, University of Michigan Medical School, 7310 CCGC, Ann Arbor, MI 48109-0946. Phone: (734) 764-2248; Fax: (734) 615-2719; E-mail: kcooney@umich.edu.

[3] The abbreviations used are: *MSR1*, macrophage scavenger receptor 1; HPC, hereditary prostate cancer; SNP, single nucleotide polymorphism; PSA, prostate-specific antigen; OR, odds ratio; FMHS, Flint Men's Health Study.

Table 1 *Rare MSR1 germ-line mutations in African-American men with prostate cancer and unaffected controls*

Mutation	Prostate cancer cases (%) (n = 134)	Unaffected men (%) (n = 340)	Fisher's exact test[a]
Ser41Tyr	1 (0.8)	0 (0)	0.278
Asp174Tyr[b]	9 (6.8)	12 (3.6)	0.143
Gly294Glu	0 (0)	0 (0)	
Pro36Ala	9 (7.0)	27 (8.1)	0.847
Arg293X	1 (0.8)	0 (0)	0.278

[a] *P*s based on two-sided test.
[b] Two unrelated cases were homozygous for the Asp174Tyr mutation.

entire sample, with the exception that age at death was used for the 29 controls that died before this date. This age variable was inserted into the models as an independent covariate. All tests were performed at the 5% significance level and using the SAS System (Cary, NC).

Haplotype-based association studies and calculation of the marker–marker linkage disequilibrium measure D′ (7) were performed using the computer program Dandelion (Green, Langefeld, and Lange, unpublished software) following the methodology described in Mohlke *et al.* (8) Briefly, a series of likelihood ratio tests were performed comparing the haplotype frequencies between cases and controls, as estimated by the expectation–maximization algorithm, for two, three, four, and five adjacent marker haplotypes (9). Statistical significance was evaluated using a permutation test based on 1000 random permutations of affection status.

Results and Discussion

The mutation frequencies for five nonsynonymous germ-line mutations are summarized in Table 1. In the first report of rare *MSR1* mutations and prostate cancer, the Asp174Tyr missense change was reported to occur with increased frequency among African-American men with apparent sporadic prostate cancer (4). In our community-based sample, the Asp174Tyr change was identified roughly twice as frequently in men with prostate cancer (6.8%) than in unaffected controls (3.6%; $P = 0.14$). In addition, 2 of 9 (22.2%) cases were homozygous for the missense change at this allele. Clinicopathologic features for the 9 cases carrying this mutation are summarized in Table 2. Among the 8 patients with available clinical data, 6 (75%) had clinically localized disease at the time of diagnosis. However, 2 cases presented with metastatic disease, with serum PSA levels of 157.8 and 1160 ng/ml, respectively.

Xu *et al.* also reported that the Asp174Tyr missense change cosegregated with prostate cancer in African-American families affected with HPC. Although formal linkage studies were beyond the scope of this case-control study, it is intriguing that at least three (37.5%) carriers of the Asp174Tyr change reported a history of prostate cancer in a first-degree relative, including one man whose family history fulfills the criteria for HPC (two brothers diagnosed with prostate cancer; Ref. 10).

Given that prostate cancer is, in general, a late-onset disease with a long asymptomatic phase, it is also notable that three of the unaffected men carrying the Asp174Tyr mutation had serum PSA levels in excess of 4 ng/ml, and at least two other unaffected carriers have a history of prostate cancer in a first-degree relative (data not shown). Moreover, the mean age of unaffected men with the Asp174Tyr change was 54.2 years, and 6 (50%) of the individuals are ≤50 years of age. This clinical data raises the possibility that, for a number of men, insufficient time may have elapsed to allow phenotypic expression of the underlying genetic variation. Indeed, misclassification of only a few controls may contribute to the lack of statistical significance for the Asp174Tyr mutation in this study sample.

The relative frequencies of the common *MSR1* sequence variants are compared for affected and unaffected men in Table 3. The relative genotype frequencies were similar for cases and controls for each of the common sequence variants with the exception of one nonsynonymous SNP in intron 5 (IVS5-59). For this SNP, heterozygosity (CA *versus* CC) was significantly more common among affected than unaffected men ($P = 0.02$). For each of the common sequence variants, the allele frequencies and age-adjusted prostate cancer ORs are summarized in Table 4. To estimate the prostate cancer risk

Table 3 *Frequencies of common MSR1 sequence variants in African-American men with prostate cancer and unaffected controls*

SNP and genotype	No. (%) of individuals with Genotype: Control subjects (n = 340)	Case subjects (n = 134)	P[a]
PRO3			
AA	125 (37.5)	54 (41.2)	0.244
AG	166 (49.9)	55 (42.0)	
GG	42 (12.6)	22 (16.8)	
INDEL1			
−/−	125 (37.7)	51 (40.2)	0.228
−/+	166 (50.0)	54 (42.5)	
+/+	41 (12.4)	22 (17.3)	
IVS5–59			
CC	329 (99.4)	122 (96.1)	0.020
CA	2 (0.6)	5 (3.9)	
AA	0 (0)	0 (0)	
P275A			
CC	287 (86.2)	118 (90.1)	0.535
CG	43 (12.9)	12 (9.2)	
GG	3 (0.9)	1 (0.8)	
INDEL7			
−/−	168 (50.9)	71 (56.8)	0.478
−/+	134 (40.6)	43 (34.4)	
+/+	28 (8.5)	11 (8.8)	

[a] χ^2 test.

Table 2 *Clinicopathologic features of nine cases with Asp174Tyr missense mutation*

Case	Genotype	Family history of prostate cancer	Age (years)	Serum PSA at diagnosis (ng/ml)	Clinical stage	Gleason Sum[a]	Type of therapy	Pathologic stage[b]	Metastatic disease
1	Homozygous	No	68.3	20.7	T1cNXMO	7	External radiation		No
2[c]	Homozygous	No	71.7	6.2	T2bNXMO	7	External radiation		No
3	Heterozygous	Yes[d]	59.7	5.2	T1cNXMO	7	Radical prostatectomy	T3aNXMX	No
4	Heterozygous	Yes[e]	63.9	157.8	T3cNXM1	9	Hormonal		Yes
5	Heterozygous	No	64.8	1160.0	T2bNXM1	8	Hormonal chemotherapy		Yes
6	Heterozygous	No	68.5	11.0	T1aNXMO	4	External radiation		No
7	Heterozygous	No	61.4	2.0	T2bNXMO	5	Radical prostatectomy	N/A	No
8	Heterozygous	N/A	57.4	0.6	T2aNXMO	6	Radical prostatectomy	T2aNOMX	No
9[f]	Heterozygous	Yes[g]	50.1	N/A	N/A	N/A	N/A	N/A	N/A

[a] Pathologic Gleason Sum is reported whenever available; otherwise, biopsy Gleason Sum is reported.
[b] For cases undergoing radical prostatectomy.
[c] Additionally carries Ser41Tyr missense mutation.
[d] Brother with prostate cancer.
[e] Two brothers with prostate cancer.
[f] Serum sample provided but epidemiologic questionnaire not completed.
[g] Father with prostate cancer.

Table 4 *Common sequence variant allele frequencies and age-adjusted OR estimates for prostate cancer among African-American cases and controls*

Allele	Allele frequencies (%)		Fisher's exact test for allele (*P*)	Age-adjusted OR (95% confidence interval)
	Control subjects	Case subjects		
PRO3 "G"	37.5	37.8	0.940	0.84[a] (0.55, 1.29)
INDEL1 "+"[b]	37.4	38.6	0.761	0.88[c] (0.57, 1.36)
IVS5–59 "A"	0.3	2.0	0.020	2.90[d] (0.50, 16.79)
P275A "C"	92.6	94.7	0.312	0.72[e] (0.37, 1.40)
INDEL7 "−"[f]	71.2	74.0	0.457	0.78[g] (0.50, 1.19)

[a] Relative odds for AG or GG genotypes *vs.* AA (referent OR = 1.00).
[b] +, the presence of the 15-bp sequence "GAATGCTTTATTGTA."
[c] Relative odds for +/− or +/+ genotypes *vs.* −/− (referent OR = 1.00).
[d] Relative odds for CA or AA genotypes *vs.* CC (referent OR = 1.00).
[e] Relative odds for CG or GG genotypes *vs.* CC (referent OR = 1.00).
[f] −, the absence of the 3-bp sequence "TTA."
[g] Relative odds for +/− or +/+ genotypes *vs.* −/− (referent OR = 1.00).

associated with each sequence variant, we compared prostate cancer risk for one genotype to the combined risk associated with two other genotypes as described previously (3). In this analysis, although the IVS5-59 variant was associated with an increased risk of prostate cancer (OR = 2.9, 95% confidence interval 0.5–16.79), this finding did not reach statistical significance. Haplotype analyses using the five common polymorphisms for all possible combinations of two, three, four, and five adjacent markers revealed no statistically significant findings (minimum $P = 0.2$ obtained for two marker haplotypes defined by IVS5-59 and P275A). Contrary to the findings of Xu *et al.* (3), evidence for marker–marker linkage disequilibrium was observed for all marker pairings, with values of D′ ranging from 1 for marker pairings PRO3—INDEL1, IVS5–59—P275A, IVS5–59—INDEL7 to 0.42 for the marker pairing INDEL1—INDEL7.

Our data provide some additional evidence for a potential link between prostate cancer and germ-line *MSR1* mutations in African-American men. Xu *et al.* (4) reported the presence of the Asp174Tyr change in 6 of 48 African-American men with non-HPC *versus* only 2 of 110 unaffected African-American men. Indeed, when our data are considered in conjunction with these findings, Asp174Tyr mutations are seen in 15 of 182 (8.2%) cases *versus* only 14 of 450 (3.1%) controls ($P < 0.05$). Furthermore, the presence of homozygosity at Asp174Tyr in 2 cases from our sample, one of whom also carried the Ser41Tyr change, raises interesting questions regarding the impact of multiple germ-line mutations on the biology and function of *MSR1,* as well as the coincident effect on prostate cancer risk.

In general, however, our data provide limited support for an association, in African-American men, between prostate cancer and the five common *MSR1* sequence variants. We evaluated each of the common variants that have been reported previously to confer increased prostate cancer risk among men of European descent (3). Statistically significant differences in allele frequencies, among cases and controls, were observed for only one (IVS5-59) of the five sequence variants (Table 3). However, the overall prevalence of this mutation (IVS5-59) in our sample was sufficiently low (2% of cases *versus* 0.3% of controls) that it may be more appropriately classified as a rare mutation rather than a common sequence variant. This discrepancy notwithstanding, after adjustment for age, none of the sequence variants was associated with a significantly increased risk of prostate cancer (Table 4). The results were similar when the control sample was limited to those men who were >50 years of age with screening PSA value(s) < 4 ng/ml (data not shown). Thus, for the common *MSR1* sequence variants, with the possible exception of IVS5-59, our results in a sample of African-American men are inconsistent with those described previously (3).

Xu *et al.* (3) reported previously that each of the common *MSR1* sequence variants, with the exception of INDEL7, was associated with an elevated risk for prostate cancer. However, a recognized limitation of this study was the potential for population stratification, whereby the observed differences in genotype frequencies may partially reflect differing genetic backgrounds among case and control subjects. In contrast, it is more likely that men in our community-based sample come from similar genetic backgrounds, thereby minimizing population stratification and potentially explaining the lack of an association, in our sample, between common *MSR1* sequence variants and prostate cancer risk. Furthermore, it is important to recognize that the study by Xu *et al.* (3) included only men of European descent, whereas our sample was comprised exclusively of African-American men. As a result, it is reasonable that a different conclusion may be reached for African-American men without necessarily compromising the validity and importance of this association in Caucasian men.

There are several limitations to our study. First, we recognize that the relatively small sample size may result in low statistical power for some of our analyses. In addition, selection bias is a potential threat to the validity of all observational studies. Among control subjects in FMHS, <60% of men that completed the initial epidemiological interview participated in the blood draw and clinical examination components of the study. Factors associated with participation in the clinical phases of the project include young age, a family history of prostate cancer, and the presence of urological symptoms (11). Although nonresponse bias is a concern for epidemiological studies of behavioral risk factors, we have no reason to believe that participants and nonparticipants differ systematically with respect to their genetic background.

In conclusion, our analysis of *MSR1* variants in 474 African American men from a community-based study of prostate cancer provides some additional support for an association between rare germ-line *MSR1* mutations and prostate cancer risk. Specifically, we observed that the Asp174Tyr missense mutation is found nearly twice as frequently among prostate cancer cases compared with controls. Although this difference in mutation frequency did not reach statistical significance in our sample, our findings are nonetheless consistent with the hypothesis that this, and potentially other, rare germ-line mutations may mediate prostate cancer risk among African-American men (4). In addition, the IVS5-59 sequence variant may also modify prostate cancer risk among African-American men, and further investigation into the prevalence and functional significance of this change is warranted. We were unable to demonstrate, in African-American men, an association between four other *MSR1* common sequence variants and prostate cancer risk. This study adds to an expanding body of epidemiological evidence in support of the hypothesis that germ-line *MSR1* mutations are risk factors for prostate cancer. Although the evidence from our study is admittedly modest, the public health burden of prostate cancer in the African-American community warrants further investigation of this potential genetic risk factor.

Acknowledgments

We thank all of the participants in the FMHS, Hong Chen for organization and quality assurance of DNA samples, Kay Doerr and Julie McLaughlin for coordinating FMHS samples and data, and Dr. David Schottenfeld for reviewing this manuscript.

References

1. Xu, J., Zheng, S. L., Hawkins, G. A., Faith, D. A., Kelly, B., Isaacs, S. D., Wiley, K. E., Chang, B., Ewing, C. M., Bujnovszky, P., Carpten, J. D., Bleecker, E. R., Walsh, P. C., Trent, J. M., Meyers, D. A., and Isaacs, W. B. Linkage and association

studies of prostate cancer susceptibility: evidence for linkage at 8p22–23. Am. J. Hum. Genet., *69:* 341–350, 2001.
2. Gibbs, M., Stanford, J. L., Jarvik, G. P., Janer, M., Badzioch, M., Peters, M. A., Goode, E. L., Kolb, S., Chakrabarti, L., Shook, M., Basom, R., Ostrander, E. A., and Hood, L. A genomic scan of families with prostate cancer identifies multiple regions of interest. Am. J. Hum. Genet., *67:* 100–109, 2000.
3. Xu, J., Zheng, S. L., Komiya, A., Mychaleckyj, J. C., Isaacs, S. D., Chang, B., Turner, A. R., Ewing, C. M., Wiley, K. E., Hawkins, G. A., Bleecker, E. R., Walsh, P. C., Meyers, D. A., and Isaacs, W. B. Common sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk. Am. J. Hum. Genet., *72:* 208–212, 2003.
4. Xu, J. F., Zheng, S. L., Komiya, A., Mychaleckyj, J. C., Isaacs, S. D., Hu, J. J., Sterling, D., Lange, E. M., Hawkins, G. A., Turner, A., Ewing, C. M., Faith, D. A., Johnson, J. R., Suzuki, H., Bujnovszky, P., Wiley, K. E., DeMarzo, A. M., Bova, G. S., Chang, B. L., Hall, M. C., McCullough, D. L., Partin, A. W., Kassabian, V. S., Carpten, J. D., Bailey-Wilson, J. E., and Isaacs, W. B. Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk. Nat. Genet., *32:* 321–325, 2002.
5. Stanford, J. L., Stephenson, R. A., Coyle, L. M., Cerhan, J., Correa, R., Eley, J. W., Gilliland, F., Hankey, B., Kolonel, L. N., Kosary, C., Ross, R., Severson, R., and West, D. Prostate Cancer Trends 1973–1995, SEER Program, National Cancer Institute. NIH Pub. No. 99-4543. Bethesda, MD, 1999.
6. Cooney, K. A., Strawderman, M. S., Wojno, K. J., Doerr, K. M., Taylor, A., Alcser, K. H., Heeringa, S. G., Taylor, J. M., Wei, J. T., Montie, J. E., and Schottenfeld, D. Age-specific distribution of serum prostate-specific antigen in a community-based study of African-American men. Urology, *57:* 91–96, 2001.
7. Excoffier, L., and Slatkin, M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol. Biol. Evol., *12:* 921–927, 1995.
8. Mohlke, K. L., Lange, E. M., Valle, T. T., Ghosh, S., Magnuson, V. L., Silander, K., Watanabe, R. M., Chines, P. S., Bergman, R. N., Tuomilehto, J., Collins, F. S., and Boehnke, M. Linkage disequilibrium between microsatellite markers extends beyond 1 cM on chromosome 20 in Finns. Genome Res., *11:* 1221–1226, 2001.
9. Lewontin, R. C. On measures of gametic disequilibrium. Genetics, *120:* 849–852, 1988.
10. Carter, B. S., Bova, G. S., Beaty, T. H., Steinberg, G. D., Childs, B., Isaacs, W. B., and Walsh, P. C. Hereditary prostate cancer: epidemiologic and clinical features. J. Urol., *150:* 797–802, 1993.
11. Heeringa, S. G., Alcser, K. H., Doerr, K., Strawderman, M., Cooney, K., Medbery, B., and Schottenfeld, D. Potential selection bias in a community-based study of PSA levels in African-American men. J. Clin. Epidemiol., *54:* 142–148, 2001.

INS PstI Polymorphism and Prostate Cancer in African-American Men

Gina B. Claeys,[1] Aruna V. Sarma,[2,3] Rodney L. Dunn,[2] Kimberly A. Zuhlke,[1] Jennifer Beebe-Dimmer,[2,3] James E. Montie,[2] Kirk J. Wojno,[2] David Schottenfeld,[1,3] and Kathleen A. Cooney[1,2]*

[1]*Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan*
[2]*Department of Urology, University of Michigan Medical School, Ann Arbor, Michigan*
[3]*Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, Michigan*

BACKGROUND. Both prostate cancer and diabetes mellitus are common diseases in African-American men. High insulin levels and insulin resistance have been implicated in prostate cancer development, which has prompted a recent investigation of a possible role for germline variation in the insulin gene (*INS*) and prostate cancer risk.
METHODS. Four hundred sixty-six African-American men with and without prostate cancer from the Flint Men's Health Study were typed for the *INS Pst1* genotype using restriction digest and direct sequencing. An association between the *Pst1* genotype and prostate cancer was examined using crude and age-adjusted logistic regression models.
RESULTS. African-American men who were homozygous for the *INS PstI* CC genotype were 1.59 times more likely to be diagnosed with prostate cancer compared to men with the TT or TC genotypes (95% CI = 0.93–2.72). The association appeared stronger among diabetics compared to non-diabetics; however this observation was not statistically significant.
CONCLUSIONS. Our study, taken together with the report of [Ho et al. Br J Cancer 88:263–269, 2003], suggests that the *INS Pst1* CC genotype is associated with prostate cancer risk in African-American men. Germline variation in the *INS* gene should be more fully explored in multiethnic studies to elucidate the molecular variant(s) associated with prostate carcinogenesis. *Prostate 65: 83–87, 2005.* © 2005 Wiley-Liss, Inc.

KEY WORDS: *INS*; **polymorphism; African-American; prostate cancer**

INTRODUCTION

Prostate cancer is the most common cancer among American men and is the second leading cause of cancer deaths in the United States. Age, African-American race, and family history are well-established risk factors for this disease [1]. African-American men have an approximately 1.5-fold higher incidence of prostate cancer and a nearly 2.5-fold higher incidence of distant disease compared to Caucasian men [2]. The likely causes which explain the racial differences in prostate cancer development and progression have not been fully elucidated, but likely include genetic, environment, and sociological factors.

African-American men are also more commonly diagnosed with type 2 diabetes mellitus compared to Caucasian men. Using data from the Third National Health and Nutrition Examination Survey (NHANES), Robbins et al. [3] showed that African-American race is associated with an approximately 50% increase in the risk of type 2 diabetes mellitus in men (age-adjusted OR = 1.43, 95% CI = 1.03–1.99). In this study, the excess risk of diabetes could not be explained by other factors including socioeconomic variables and body size. The

Grant sponsor: S.P.O.R.E. in Prostate Cancer; Grant number: P50CA69568.

*Correspondence to: Kathleen A. Cooney, MD, 7310 CCGC 0946, 1500 East Medical Center Drive, Ann Arbor, MI 48109-0946.
E-mail: kcooney@umich.edu

Received 8 December 2004; Accepted 7 March 2005
DOI 10.1002/pros.20271
Published online 2 May 2005 in Wiley InterScience (www.interscience.wiley.com).

proposed pathophysiological mechanism for type 2 diabetes is peripheral insulin resistance in combination with relative insulin deficiency.

Several groups have explored the potential relationship between prostate cancer and various aspects of insulin physiology. Using data from a case-control study of prostate cancer conducted in China, Hsing et al. [4] demonstrated that men in the highest tertile of insulin resistance had a 2.8-fold increased risk or prostate cancer (95% CI = 1.6–4.7). Other reports have shown an increased risk of prostate cancer in men with high insulin levels [5] as well as a relationship between serum insulin levels and advanced tumor stage and risk of prostate cancer recurrence [6,7].

Given the potentially important role of insulin in prostate cancer development, Ho et al. [8] tested genetic variation in the insulin gene (*INS*) to determine its potential association with prostate cancer susceptibility. Tight linkage disequilibrium has previously been reported between 10 SNPs spanning a 4.1 kb genomic segment on chromosome 11p15.5 containing the *INS* gene, therefore these investigators elected to test a single SNP (+1127 *INS* Pst1) in the 3′ untranslated region (UTR) of the gene. Using DNA samples from a hospital-based case-control study, Ho et al. [8] detected an increased risk of prostate cancer in men who had the homozygous CC *INS* Pst1 genotype (OR = 1.74, 95% CI = 0.99–3.05). The strongest association, however, was observed in the subset of men who were over age 54 years, either African-American or Caucasian, and who did not report a personal history of diabetes (OR = 6.33, 95% CI 1.87–21.4).

The Flint Men's Health Study (FMHS) is a community-based case-control study of prostate cancer in African-American men between the ages of 40–79. To further explore the potential association between diabetes, germline variation in the *INS* gene and prostate cancer, we genotyped 466 men with and without prostate cancer for the +1127*INS* Pst1 SNP and herein describe the association of the CC genotype with prostate cancer in our study population.

MATERIALS AND METHODS

Control Subjects

Data collection for the FMHS began in 1996 and concluded in 2002. Informed consent was obtained from all study participants, and protocols were approved by the Institutional Review Board at the University of Michigan Medical School. As described previously, disease-free controls were identified from a probability sample of African-American men in Flint, Michigan or in selected census tracts in neighboring Beecher Township (Genesee County, Michigan) [9]. The source population for controls includes all African-American men aged 40–79 living in Genesee County. Men in older age groups were over-sampled to increase the number of eligible subjects for analyses pertaining to prostate cancer. From the initial sample of 943 men, 732 were willing to participate and determined to be eligible to complete the detailed in-home epidemiologic interview, which covered information on health behaviors such as smoking, drinking and physical activity; occupational exposures, general health condition and medical history of chronic illnesses, family history of prostate cancer and demographic information.

Subjects were then asked to undergo a prostate cancer screening protocol, which included providing a blood sample for a serum total prostate-specific antigen (PSA) measurement, and undergoing a comprehensive urological examination. A total of 379 men completed the interview, blood draw and clinical examination. Men with an abnormal digital rectal examination and/or elevated total PSA concentration ($\geq$4.0 ng/ml) were referred for prostate biopsy. Twenty eight men subsequently diagnosed with biopsy-confirmed prostate cancer were included in the study as cases; the remaining 351 were included in our control sample.

Case Subjects

Prostate cancer case recruitment from the same community began in 1999 and was completed in 2002. Cases were identified using the Genesee County Community-Wide Hospital Oncology Program (CHOP) registry, which includes the three hospitals for the county: Hurley Hospital, Genesys Regional Medical Center, and McLaren Regional Medical Center. Case participation in the study required (1) an in-home interview as in the controls; (2) a review of the hospital and registry records for information on stage, Gleason's grade of differentiation, treatment, and pre-diagnosis PSA value; (3) anthropometric measurements; and (4) a blood sample for DNA and freezer storage of serum and plasma. An expert in genitourinary pathology at the University of Michigan reviewed the pathological material and assigned Gleason grade. One hundred twenty five African-American men aged 40–79 who were residents of Genesee County and who had been diagnosed with prostate cancer completed all aspects of the case protocol. The final case group includes the 28 newly diagnosed cases identified during recruitment of potential controls.

Laboratory Methods

Genomic DNA was extracted from whole blood using a Puregene DNA extraction kit (Gentra Systems, Inc., Research Triangle Park, NC). Of the 351 control and 125 case samples, 9 and 1, respectively, were

removed due to technical reasons. The *INS PstI* genotypes were analyzed for 342 controls without the clinical diagnosis of prostate cancer and 124 cases. Genotyping of the *INS PstI* was performed using minor modifications of the restriction fragment length polymorphism (RFLP) analysis described by Ho et al. [8]. Primers were designed to amplify a 503 bp region of the *INS* gene containing the +1127 *Pst1* SNP. The forward primer sequence was 5′CGGGGGAAGGAGGTGGGACAT, while the reverse primer sequence was 5′ACAACAGTGCCGGGAAGTGGGG. PCR primers were purchased from Invitrogen (Carlsbad, CA). All genotypes were independently interpreted by a second researcher, and 18% of genotypes were confirmed by repeat restriction digest. Direct sequencing was used to independently confirm 7% of all samples, and complete concordance was observed between the genotypes obtained by sequencing and RFLP analysis.

Statistical Analysis

The distribution of the *INS PstI* genotype along with other sociodemographic and clinical characteristics were examined in the overall study population as well as by case status. Differences in genotype frequency between cases and controls were tested using the Mantel–Haenszel Chi-square test. Unconditional logistic regression was performed to obtain multivariable odds ratios with associated 95% confidence intervals for the association between *INS PstI* genotype and prostate cancer, adjusting for potentially important covariates and prostate cancer. The TC and TT genotypes were combined in the final analysis given the rarity of the TT genotype. The final models adjusted simultaneously for *INS PstI* genotype, age, BMI, and diabetes. Differences in the associations between *INS PstI* genotype and prostate cancer by BMI and diabetes status were tested using the Breslow–Day test for homogeneity.

Approximately 50% of FMHS control men were tested for prostate cancer at multiple time points. Therefore, to avoid lead time bias in the multivariable analysis, age was calculated based on the same date for all cases and controls. This date was the most recent follow-up date from the entire sample, with the exception that age at death was used for the 37 controls that died prior this date. Subjects were considered to have a positive family history of prostate cancer if they reported that a father, son, and/or brother had ever been diagnosed with prostate cancer. Subjects were considered to have diabetes based on the subjects' self-report of physician-diagnosed diabetes. BMI was calculated by dividing weight in kilograms by height in meter squared, using measurements of weight and height obtained during the clinical exam and these measures were then categorized based on the World Health Organization (WHO) definitions of overweight and obesity: obese ($\geq$30), overweight ($25 \leq \text{BMI} < 30$), or normal BMI (<25). The Cochran–Armitage trend test was calculated to determine if variables with more than two categories had a linear trend of increasing risk of prostate cancer for each successive category. All analyses were performed using the Statistical Analysis System (SAS v. 8.2, Cary, NC). Two tailed tests were used for all comparisons and *P*-values of less than 0.05 were considered statistically significant.

RESULTS

Overall, the men with prostate cancer that were genotyped in this report were significantly older than men without prostate cancer, 64.3 years versus 56.1 years ($P < 0.0001$). A positive family history for prostate cancer was reported by 21.8% of cases compared to 16.7% of controls ($P = 0.22$). Both BMI and diabetes were positively associated with prostate cancer; however, these associations were not statistically significant (Table I).

TABLE I. Distribution of Clinical Characteristics by Case/Control Status and Associated Odds Ratio for the Diagnosis of Prostate Cancer

	Prostate cancer cases (n = 124)	Disease-free controls (n = 342)	Odds ratio (95% CI)	*P*-value
BMI				0.126
Normal (<25)	25.4%	33.3%	—	
Overweight (25–29.9)	40.2%	36.3%	1.45 (0.87–2.44)	
Obese (≥30)	34.4%	30.4%	1.49 (0.87–2.54)	
Diabetes				0.286
Yes	22.6%	17.8%	1.34 (0.81–2.22)	
No	77.4%	82.2%	—	

The *INS PstI* genotypes obtained from 466 FMHS participants exhibited Hardy–Weinberg equilibrium. The CC genotype was most frequently observed in both cases and controls, 83.9% and 76.6%, respectively (Table II). The TT genotype was observed in only three men or 0.6% of the entire study sample. Our results suggest that men with the *INS PstI* CC genotype had a 1.59-fold greater risk of prostate cancer diagnosis as compared to men with the TT or TC genotype (95% CI = 0.93–2.72). We also pooled the FMHS data with the genotype data from 96 African-American men with prostate cancer and 67 disease-free controls reported by Ho et al. [8]. The pooled estimate of prostate cancer risk associated with the *INS Pst1* CC genotype was 1.53 (95% CI = 1.01–2.32). Among the cases, there was no significant association between the *INS PstI* CC genotype and prostate cancer grade (comparing Gleason score ≤6 to Gleason score ≥7 cases) or age at diagnosis (data not shown).

No differences were observed in the risk for prostate cancer associated with the *INS PstI* CC genotype in obese compared to non-obese men (Table III). The risk of prostate cancer associated with the CC genotype appeared greater among men who reported the diagnosis of diabetes (OR = 2.35, 95% CI = 0.76–7.26) compared to those who did not (OR = 1.55, 95% CI = 0.82–2.94), although the difference was not significant.

TABLE III. Age-Adjusted Odds Ratios of *INS PstI* CC Genotype by BMI and Diabetes Strata

	Odds ratio (95% CI) CC versus TC/TT
BMI	
Obese (≥30)	1.67 (0.62–4.51)
Non-obese (<30)	1.75 (0.89–3.45)
Diabetes	
Yes	2.35 (0.76–7.26)
No	1.55 (0.82–2.94)

DISCUSSION

This is the first independent confirmation of the association between the *INS PstI* CC genotype and prostate cancer as initially described in a multiethnic, hospital-based case-control study at Albert Einstein College of Medicine in New York City. Since our study population was comprised solely of African-American men, we elected to combine the data from 466 African-American FMHS participants with the subset of 163 African-American men from the New York City study. In this combined sample of 629 men, the *INS PstI* CC genotype was associated with a 1.53-fold increase risk of prostate cancer (95% CI = 1.01–2.32).

The *INS* gene, comprised of three exons, is located between the insulin-like growth factor 2 (*IGF2*) gene and the tyrosine hydroxylase (*TH*) gene in chromosome 11p15. The *Pst1* polymorphism is located in the 3′ UTR *INS*. Given the likely role of the 3′ UTR in mRNA stability, it is possible that this polymorphism is playing a direct role in altering insulin physiology which ultimately contributes to prostate carcinogenesis. However, tight linkage disequilibrium has also been observed in the region of the *INS* gene [10] and so it is also possible that the *Pst1* C allele is linked to the disease-associated haplotype. In this case, further investigations must be completed to define the specific alteration which leads to prostate cancer in men.

Despite the fact that both our study as well as the Ho et al. [8] observed similar increases in risk of prostate cancer associated with the *Pst1 INS* C allele, there are some differences in our study designs and results that should be noted. The cases and controls in the

TABLE II. Distribution of *INS PstI* Genotypes and Other Characteristics by Case/Control Status and Associated Odds Ratios for Prostate Cancer Diagnosis With 95% Confidence Intervals

Data source	*INS PstI* genotype	Prostate cancer cases	Disease-free controls	Odds ratio (95% CI)
FMHS data				
	CC	104 (83.9%)	262 (76.6%)	1.59 (0.93–2.72)[a]
	TC	19 (15.3%)	78 (22.8%)	—
	TT	1 (0.8%)	2 (0.6%)	—
Ho et al. [8]				
African-American men	CC	78 (81.3%)	48 (71.6%)	1.72 (0.82–3.59)[a]
	TC	17 (17.7%)	18 (26.9%)	—
	TT	1 (1.0%)	1 (1.5%)	—
Pooled data				
	CC	182 (82.7%)	310 (76.8%)	1.53 (1.01–2.32)
	TC or TT	38 (17.3%)	99 (24.2%)	—

[a]Odds ratio compares CC genotype to TC and TT genotypes combined.

New York City study were identified in a clinic setting and were matched by birth year, race and country of origin. The FMHS is a community-based study in which controls were identified using probability sampling minimizing the potential for selection bias. However, one possible limitation of our study was that only approximately 50% of interviewed subjects ultimately completed all aspects of the clinical examination for prostate cancer in the first wave of the study, which may introduce bias due to non-participation. Heeringa et al. [11] examined the potential selection bias due to non-participation that may have occurred in the FMHS and observed that men who completed all aspects of the study protocol were younger, more likely to be experiencing urological symptoms and were more likely to have a family history or prostate cancer compared to those men who did not undergo clinical examination. The high rate of prostatectomy among the FMHS cases (data not shown) suggests that the FMHS cases may have earlier stage disease compared to the target population. Because prostate cancer is a late-onset disease, both our study and that of Ho et al. [8] may have enrolled men as controls who will ultimately be diagnosed with prostate cancer. However, in case-control studies, the selection of controls from the same or similar reference population assumes that these individuals to some unknown degree may later be diagnosed as cases.

The Ho et al. [8] study revealed that the strongest association between *INS* genotype and prostate cancer was in non-diabetic men who were over 55 years of age. The men who were homozygous CC also were more likely to be older and have a moderately to well-differentiated cancer (Gleason score <7). In contrast, no relationship between *INS* genotype and prostate cancer age at diagnosis or Gleason score was detected in the FMHS, and the OR attributed to the CC genotype was higher among diabetics in the FMHS. As the majority of FMHS cases presented with earlier stage disease, it is possible there was not enough variation in our sample for us to detect an association between *INS* genotype and prostate cancer severity. Racial differences is the prevalence of exposure could account for the discrepancies observed by diabetes status. Larger studies of multiethnic cohorts are necessary to further elucidate these potential relationships.

In conclusion, members of the insulin gene family have become intriguing candidates for genetic studies of prostate cancer susceptibility. Although the relationships between diabetes, insulin levels, insulin resistance, and obesity are complex, molecular studies performed in the context of carefully designed epidemiological studies have the opportunity to tease out potential associations. Our study provides some confirmation to the investigation conducted by Ho et al. [8] in that the observed risk associated with the *Pst1 INS* CC genotype in African-American men was similar to the results from their investigation conducted in a multiethnic population. More thorough studies of germline variants in this region will uncover the true causative mutation that may play a role in prostate carcinogenesis directly or indirectly through insulin action.

ACKNOWLEDGMENTS

We would like to thank all of the participants in the FMHS. Furthermore, the efforts of Kay Doerr, Julie McLaughlin, and Stephen Gruber have been invaluable toward the completion of this study.

REFERENCES

1. Brawley OW, Knopf K, Thompson I. The epidemiology of prostate cancer part II: The risk factors. Semin Urol Oncol 1998;16:193–201.
2. Sarma AV, Schottenfeld D. Prostate cancer incidence, mortality, and survival trends in the United States: 1981–2001. Semin Urol Oncol 2002;20:3–9.
3. Robbins JM, Vaccarino V, Zhang HP, Kasl SV. Excess type 2 diabetes in African-American women and men aged 40–74 and socioeconomic status: Evidence from the Third National Health and Nutrition Examination Survey. J Epidemiol Community Health 2000;54:839–845.
4. Hsing AW, Gao YT, Chua S, Deng J, Stanczyk FZ. Insulin resistance and prostate cancer risk. J Natl Cancer Inst 2003;95: 67–71.
5. Hsing AW, Chua S Jr, Gao YT, Gentzschein E, Chang L, Deng J, Stanczyk FZ. Prostate cancer risk and serum levels of insulin and leptin: A population-based study. J Natl Cancer Inst 2001;93: 783–789.
6. Lehrer S, Diamond EJ, Stagger S, Stone NN, Stock RG. Serum insulin level, disease stage, prostate specific antigen (PSA) and Gleason score in prostate cancer. Br J Cancer 2002;87:726–728.
7. Lehrer S, Diamond EJ, Stagger S, Stone NN, Stock RG. Increased serum insulin associated with increased risk of prostate cancer recurrence. Prostate 2002;50:1–3.
8. Ho GY, Melman A, Liu SM, Li M, Yu H, Negassa A, Burk RD, Hsing AW, Ghavamian R, Chua SC. Polymorphism of the insulin gene is associated with increased prostate cancer risk. Br J Cancer 2003;88:263–269.
9. Cooney KA, Strawderman MS, Wojno KJ, Doerr KM, Taylor A, Alcser KH, Heeringa SG, Taylor JM, Wei JT, Montie JE, Schottenfeld D. Age-specific distribution of serum prostate-specific antigen in a community-based study of African-American men. Urology 2001;57:91–96.
10. Lucassen AM, Julier C, Beressi JP, Boitard C, Froguel P, Lathrop M, Bell JI. Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. Nat Genet 1993;4:305–310.
11. Heeringa SG, Alcser KH, Doerr K, Strawderman M, Cooney K, Medbery B, Schottenfeld D. Potential selection bias in a community-based study of PSA levels in African-American men. J Clin Epidemiol 2001;54:142–148.

Prostate Cancer and Prostatic Diseases (2006) 9, 50–55
 1365-7852/06 $30.00
www.nature.com/pcan

ORIGINAL ARTICLE

Polymorphisms in the prostate-specific antigen gene promoter do not predict serum prostate-specific antigen levels in African-American men

JL Beebe-Dimmer[1,2], LA Lange[3], JE Cain[4], RC Lewis[4], AM Ray[4], AV Sarma[1,2], EM Lange[3,5] and KA Cooney[1,4]

[1]*Department of Urology, University of Michigan Medical School, Ann Arbor MI, USA;* [2]*Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor MI, USA;* [3]*Department of Genetics, University of North Carolina, Chapel Hill, NC, USA;* [4]*Department of Internal Medicine, University of Michigan Medical School, Ann Arbor MI, USA and* [5]*Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA*

A major problem with the use of serum prostate-specific antigen (PSA) in predicting prostate cancer risk is the considerable variability of such measurements. Cramer *et al*. identified a set of single-nucleotide polymorphisms (SNPs) in the upstream regulatory region of the PSA gene that were each associated with increased promoter activity and serum PSA, further suggesting that genotyping these SNPs could be useful in improving the predictive value of PSA screening. In order to replicate this finding, DNA samples from 475 African-American men were genotyped for the same SNPs and no association was observed with either serum PSA level or prostate cancer diagnosis.
Prostate Cancer and Prostatic Diseases (2006) **9,** 50–55. doi:10.1038/sj.pcan.4500840; published online 25 October 2005

Keywords: prostate-specific antigen; single-nucleotide polymorphism; PSA gene

Introduction

Prostate cancer is the most common cancer diagnosed among men in the US with an expected 232 000 new cases in 2005. [1] In addition to increasing age and family history, race is one of the most important recognized risk factors for the disease.[2] African-American men have an approximately 1.6-fold greater chance of being diagnosed with prostate cancer compared to Caucasian men and a 2.4-fold greater chance of dying from the disease.[1] There are multiple variables that likely contribute to the racial disparity in both incidence and mortality including genetic, environmental, and sociological factors.

Prostate-specific antigen (PSA) was first proposed as a serum marker for the early detection of prostate cancer nearly 29 years ago.[3] PSA is a serine protease that is expressed in both normal prostate epithelium as well as prostate cancer. Although serum PSA elevations greater than 4.0 ng/ml are associated with prostate cancer, the positive predictive value (PPV) of a serum PSA value between 4.0 and 10 ng/ml is only 20–30%, although the PPV rises to 40–70% for serum PSA values greater than 10 ng/ml.[4] A recent analysis of data from the Prostate Cancer Prevention Trial (PCPT) has demonstrated that prostate cancer can also be diagnosed in men with PSA values generally considered to be in the normal range. In the PCPT, approximately 15% of men with normal digital rectal examinations and PSA values consistently under 4.0 ng/ml over a period of 7 years were determined to have prostate cancer on biopsies performed at the end of the study.[5] Clearly, more research needs to be done to improve the use of serum PSA as a screening test for prostate cancer in asymptomatic men.

Transcriptional regulation of the PSA gene is mediated through binding of the androgen receptor to regions of the promoter containing androgen response elements (AREs).[6] In 2003, Cramer *et al*.[7] reported that specific germline genetic polymorphisms in the promoter region of the PSA gene were associated with higher serum PSA levels among 409 Caucasian male subjects (mean age 63.7 years). They further suggested that these polymorphisms may be useful in refining the recommendations for use of serum PSA levels for prostate cancer screening. Given the importance of prostate cancer in African-American men, we set out to explore the potential association between selected functional single-nucleotide

Correspondence: Dr JL Beebe-Dimmer, 109 S. Observatory Street RM 2004, School of Public Health, Ann Arbor, MI 48109-2029, USA.
E-mail: jbeebe@med.umich.edu
Received 12 July 2005; accepted 24 August 2005; published online 25 October 2005

polymorphisms (SNPs) in the PSA gene with serum PSA levels and prostate cancer risk using samples from the Flint Men's Health Study.

Materials and methods

The Flint Men's Health Study (FMHS) is a community-based study of prostate cancer in African-American men between the ages of 40 and 79 years. In 1996, 730 men were recruited to participate in the study from a probability sample residing in the city of Flint and surrounding communities in Genesee County, Michigan. Subjects completed a detailed in-home interview that collected information on sociodemographics, potential risk factors for prostate cancer, and a complete medical history. Subjects were also asked to participate in a clinical examination which included measurement of serum PSA (free and total) and a comprehensive urologic examination. Exam participants were asked to refrain from sexual activity for at least 24 h prior to the blood draw and all venipuncture was performed before referral for digital rectal exam as both may cause transitory increases in PSA level. Men with an elevated total PSA ($\geqslant$4.0 ng/ml) or an abnormal digital rectal exam were referred for prostate biopsy. Of the 730 men who completed the initial interview, 379 participated in the clinical exam. A total of 10 subjects were diagnosed with prostate cancer as a consequence of the protocol which resulted in a final control sample of 369 men. Attempts were made to follow the study participants and in the 5 years after control recruitment, an additional 18 control men were diagnosed with prostate cancer. Four controls with a total PSA level greater than 9.0 ng/ml were also excluded from the analysis following the analysis plan of Cramer *et al*.[7] A sufficient DNA sample was available for genotyping on 339 of the remaining controls.

Cases were recruited from the same community from 1999 to July 2002. Men who were between the ages of 40 and 79 years at the time of prostate cancer diagnosis (between 1995 and 2002) were eligible to participate in the study. Cases also completed a detailed epidemiologic interview and provided a blood sample. Medical records were reviewed to extract information related to prostate cancer diagnosis including clinical and pathologic stage, Gleason grade, prediagnostic PSA, and treatment. A total of 136 cases were ultimately recruited to participate in the study. Informed consent was obtained from all study participants and the research protocol has been approved by the Institutional Review Board of the University of Michigan. For both cases and controls, genomic DNA was isolated from whole blood using the Puregene DNA Purification Kit (Gentra Systems, Minneapolis, MN, USA).

SNP selection and genotyping methods

Three SNPs in the 5′ promoter region of the PSA gene were selected: –4643 A/G (rs925013), –5412 T/C (rs2739448), and –5429 T/G (rs2569733). Each of these SNPs were present at greater than 5% frequency in a Caucasian sample and also were shown to be associated with PSA levels.[7] For the –4643 A/G (rs925013) SNP, Assays on Demand SNP genotyping products (Applied Biosystems, Foster City, CA, USA) were used. These consist of a 20 $\times$ mixture of PCR primers and dye labeled TaqMan MGB probes designed to genotype a specific SNP within a given sequence. PCR reactions were performed in a 384-well plate format with 2.25 μl of genomic DNA (4 ng/μl), 0.25 μl 20 $\times$ SNP Genotyping Assay Mix, 1.0 μl 5 $\times$ Real Time Ready Mastermix (Qbiogene, Montreal, Canada), and 1.5 μl dH_20 for a total reaction volume of 5.0 μl. For the remaining two SNPSs, Assays by Design SNP genotyping products (Applied Biosystems, Foster City, CA, USA) were employed. These assays were custom-designed and produced based upon target sequence submission by using Assay by Design File Builder (Applied Biosystems, Foster City, CA, USA). PCR reactions were performed in a 384-well plate format with 2.375 μl of genomic DNA (4 ng/μl), 0.125 μl 40 $\times$ SNP Genotyping Assay Mix, 1.0 μl 5 $\times$ Real Time Ready Mastermix (Qbiogene, Montreal, Canada), and 1.5 μl dH_20 for a total reaction volume of 5.0 μl. Assays were optimized to use a universal thermal cycling protocol with an initial hold at 95°C for 10 min followed by 40 cycles at 92°C for 15 s and a combined annealing extension step of 60°C for 1 min. Allelic discrimination was then performed on ABI Prism 7900 HT Sequence Detection System and analyzed using SDS 2.1 software (Applied Biosystems, Foster City, CA, USA). At least 23% of the sample set was duplicated on this platform for verification of results.

Validation of results

Validation by restriction enzyme digest was performed on 15% of the sample set for the –4643 A/G (rs925013) SNP. PCR primers (5′ CAGGGATTATCTTCAGCACTTA-CA 3′, 5′ACTGGCCAGCTGGGAATAGAGATA 3′) were designed using Primer Select software (DNASTAR Inc., Madison, WI, USA) and purchased from Invitrogen Life Technologies (Carlsbad, CA, USA). Each reaction contained 5.0 μl 10 $\times$ PCR buffer, 1.0 μl 50 mM $MgCl_2$, 1.0 μl 10 mM dNTPs, 5.0 μl each of the two PCR primers at 5 μM, 2.0 μl genomic DNA at 20 ng/μl, 0.5 μl Platinum *Taq* Polymerase (Invitrogen Life Technologies), and 30.5 μl ddH_20 for a total reaction volume of 50.0 μl. PCR products were digested using 40.0 μl PCR product, 2.0 μl *Nco*I restriction enzyme, 5.0 μl 10 $\times$ NE Buffer #4 (New England Biolabs Inc., Beverly, MA, USA), and 3.0 μl ddH_2O for a total reaction volume of 50.0 μl. Samples were placed in a 37°C water bath and allowed to incubate for at least 4 h. Digested products were analyzed on 3% agarose gels.

Validation by direct sequencing was performed on 14% of the sample set for the 5412 T/C (rs2739448), and –5429 T/G (rs2569733) SNPs. For direct sequence analysis, PCR products were purified using Montage PCR Centrifugal Filter Devices (Millipore, Billerica, MA, USA). Cycle sequencing was performed in the forward direction using Big Dye Terminator Chemistries (Applied Biosystems, Foster City, CA, USA). Cycle sequencing reactions were purified using Performa DTR Gel Filtration Cartridges (Edge Biosystems, Gaithersburg, MD, USA) and were sequenced using an ABI 3100 Genetic Analyzer.

Statistical analysis

For each SNP, the observed genotype distribution was tested for consistency with Hardy–Weinberg equilibrium expectations using Pearson's χ^2 test (SAS version 8.2, Cary NC, USA). Lewontin's D' statistic was calculated to

estimate the strength of linkage disequilibrium between all possible pair-wise combinations of SNPs using the computer software GOLD (www.sph.umich.edu/csg/abecasis/GOLD/index.html).[8] PSA levels in our population were not normally distributed after adjustment for age at clinical exam and genotype. Similar to Cramer *et al*., a log base-10 (log10) transformation was performed on the PSA levels in order to satisfy distributional assumptions. Analysis of variance was performed to evaluate differences in mean values for log10 total PSA, log10-free PSA, and ratio of free to total PSA levels by genotype for each SNP adjusting for age. Since homozygosity for the least prevalent allele was rare in our study population (<1%), the reported *P*-values were calculated based on tests comparing the outcome measurements for heterozygotes and homozygotes for the least prevalent allele to homozygotes for the most prevalent allele for each SNP. The least-square means for log10 total, log10-free PSA, and the ratio of free to total PSA were calculated for each genotype. Logistic regression models were used to test whether each SNP was associated with prostate cancer. Finally, haplotype frequencies, the association between haplotypes and the PSA outcome variables adjusted for age, and the association between haplotypes and prostate cancer were evaluated using the program Haplo.Stat.[9]

Results

The genotype distributions for each of the three SNPs were found to be consistent with expected Hardy–Weinberg equilibrium proportions in both the case and control samples (data not shown). The observed genotype frequencies of the selected PSA promoter SNPs among the FMHS 339 control and 136 case participants are reported in Table 1. Men who were homozygous for the less prevalent allele were rare in this study population. All three SNPs were in near perfect linkage disequilibrium ($P<0.0001$ for all pairs), with pair-wise D' estimates of 0.95, 0.97, and 0.99 for the pairings (–4643, –5412), (–4643, –5429) and (–5412, –5429), respectively.

No significant differences in the frequencies of each of the three SNPs were observed between prostate cancer cases and controls (Table 2). Just one prostate cancer case (0.7%) and one control individual (0.3%) were determined to be homozygous for the least prevalent allele for all three SNPs. Approximately 15–16 % of cases and 18–19% of controls were heterozygotes for the three SNPs. Consistent with the analyses of PSA levels, statistical significance estimates for the case/control analyses (Table 2) reflect the results of single degree-of-freedom tests comparing the odds of being a case for individuals heterozygous or homozygous for the least prevalent allele to individuals with the most frequent allele. The case/control haplotype analyses failed to detect significant differences in haplotype frequencies between cases and controls (global score statistic $=0.81$, $P=0.66$; data not shown).

Table 1 Frequency of three SNPs in the PSA gene promoter region among African-American men

Location of SNP	*Nucleotide change*	*% Control chromosomes carrying variant* (n = 2 × 339)	*% Case chromosomes carrying variant* (n = 2 × 136)
–4643:rs925013 *(NcoI)*	A to G	9.3 (63/678)	8.8 (24/272)
–5412:rs2739448*(BstUII)*	T to C	9.3 (63/678)	8.8 (24/272)
–5429:rs2569733 *(FokI)*	T to G	9.9 (67/678)	8.5 (23/272)

Table 2 Frequencies of selected SNPs in the promoter region of the PSA gene among 136 men with prostate cancer and 339 healthy control men participating in the FMHS

SNP/genotype	*Prostate cancer cases* N *(%)*	*Controls*[a] N *(%)*	*OR*[b] *(95% CI)*
–4643:rs925013 *(NcoI)*			
AA	113 (83.1)	277 (81.7)	
AG	22 (16.2)	61 (18.0)	
GG	1 (0.7)	1 (0.3)	0.91 (0.54–1.55)
–5412:rs2739448 *(Bst*UI*)*			
TT	113 (83.1)	277 (81.7)	
TC	22 (16.2)	61 (18.0)	
CC	1 (0.7)	1 (0.3)	0.54 (0.47–1.55)
–5429:rs256733 *(Fok*I*)*			
TT	114 (83.8)	273 (80.5)	
TG	21 (15.4)	65 (19.2)	
GG	1(0.7)	1 (0.3)	0.80 (0.47–1.36)

[a]Controls with a serum PSA level of ⩾9.0 ng/ml were excluded.
[b]Odds ratio (OR) comparing individuals with at least one copy of the rare allele (heterozygotes or homozygotes) compared to individuals who were homozygotes for the more prevalent allele.

Age-specific PSA ranges for control subjects participating in the FMHS have been previously reported.[10] No associations were observed between different genotypes in selected SNPs in the promoter region of the PSA gene with log10 total PSA levels, log10 free PSA levels, or the ratio of free to total PSA (Table 3). Men with the GT genotype at –5429 SNP had similar total PSA levels compared to men with the TT genotype. Likewise, total PSA levels among men with the AG genotype in –5412 SNP were similar to men with the AA genotype. Since SNPs –4643 and –5412 were in almost perfect linkage disequilibrium, serum free and total PSA levels for each genotype in –4643 were nearly identical to those observed at SNP –5412. Of note, the individual who was homozygous for the rare allele at all three SNPs in this study had a higher PSA than the mean values for each of the other genotypes.

In order to examine whether specific combinations of allelic variants were associated with serum PSA levels, we performed haplotype-based association analyses (Table 4). The frequency of haplotypes –5429T/–5412T/4643A and –5429G/–5412C/4643G were approximately 90 and 9%, respectively. No significant differences in log10 total PSA, log10 free PSA, or ratio of free to total PSA levels were observed between the three selected haplotypes.

Discussion

Cramer *et al*. [7] reported that three polymorphisms (–4643 A/G, –5412 T/C, and –5429 T/G) located at the far 5′ upstream region of the promoter were significantly

Table 3 Least-square means, adjusted for age, of log10 total PSA, log10 free PSA, and the ratio of free to total PSA by genotype in PSA promoter region SNPs among 339 control participants of FMHS[a]SNP/genotype[b]

	N *(%)*	*Mean log (total PSA ng/ml) (s.d.)*	P-*value*[c]	*Mean log (free PSA ng/ml) (s.d.)*	P-*value*	*Ratio free:total PSA (s.d.)*	P-*value*
–4643:rs925013 *(Nco*I*)*							
AA	277 (81.7)	−0.039 (0.394)		−0.637 (0.345)		0.272 (0.106)	
AG	61 (18.0)	−0.048 (0.289)		−0.650 (0.230)		0.268 (0.099)	
GG	1 (0.3)	0.613 (N/A)	0.88	−0.268 (N/A)	0.98	0.132 (N/A)	0.65
–5412:rs2739448 *(Bst*UI*)*							
TT	277 (81.7)	−0.040 (0.394)		−0.636 (0.344)		0.273 (0.106)	
TC	61 (18.0)	−0.046 (0.0.294)		−0.652 (0.301)		0.265 (0.099)	
CC	1 (0.3)	0.613 (N/A)	0.86	−0.268 (N/A)	0.90	0.132 (N/A)	0.50
–5429:rs256733 *(Fok*I*)*							
TT	273 (80.5)	−0.037 (0.393)		−0.635 (0.345)		0.272 (0.105)	
TG	65 (19.2)	−0.057 (0.305)		−0.655 (0.303)		0.271 (0.104)	
GG	1 (0.3)	0.613 (N/A)	0.89	−0.268 (N/A)	0.80	0.132 (N/A)	0.84

[a]Men with a serum PSA level of $\geqslant 9.0$ ng/ml were excluded.
[b]Nucleotide position relative to the transcription start site.
[c]Age-adjusted comparison between men who are heterozygotes or homozygotes for the least prevalent allele compared to men who are homozygotes for the most prevalent allele.

Table 4 Association between PSA promoter haplotypes and log10 total PSA and log10 free PSA levels in 339 controls

	Allele				*Log10 total PSA*		*Log10 free PSA*		*PSA ratio*	
Haplotype	*-5429 T/G SNP*	*-5412 T/C SNP*	*-4643 A/G SNP*	*Estimated frequency*	*Score*[a]	P[b]	*Score*[a]	P[b]	*Score*[a]	P[b]
1	T	T	A	0.900	−0.21	0.83	0.11	0.91	0.37	0.72
2	G	C	G	0.091	0.37	0.71	−0.10	0.92	−0.71	0.46
3[c]	G	C	A	0.003	NA	NA	NA	NA	NA	NA
4[c]	G	T	A	0.003	NA	NA	NA	NA	NA	NA
5[c]	G	T	G	0.003	NA	NA	NA	NA	NA	NA
Global					0.34	0.84	0.01	0.99		

A significant *P*-value and a positive (negative) score for a particular haplotype would have suggested that the haplotype was associated with increased (decreased) levels of the PSA outcome variable.
[a]HAPLO.SCORE (http://www.mayo.edu/statgen) program statistic.
[b]χ^2 test of the score test statistic. PSA level adjusted for age.
[c]Given rarity of haplotype, haplotype not evaluated in statistical tests of association.
NA, not available.

associated with increased PSA levels using samples from Caucasian asbestos workers participating in a study of asbestos exposure and lung disease. The three SNPs were moderately prevalent in the population, with approximately 21–23% of men carrying the variant allele, and the SNPs were shown to be in strong linkage disequilibrium. Further analysis revealed that the –5429T/–5412T/–4643A haplotype, which occurred in 77% of the population, was associated with significantly lower serum PSA levels ($P<0.001$), while the –5429G/–5412C/–4643G haplotype occurring in 20% of the population was associated with higher PSA ($P=0.009$). In our study of a cohort of 339 African-American men between the ages of 40 and 79 years with PSA levels less than 9 ng/ml and no known diagnosis of prostate cancer, one haplotype (–5429T/–5412T/–4643A) was detected in 90% of the population and no relationship was detected between any haplotype and PSA level.[7]

A number of studies have examined associations between genetic polymorphisms in the PSA gene, PSA levels, and prostate cancer risk. The PSA gene is a member of a family of 15 kallikrien genes clustered on chromosome 19q13.3–13.4.[11] Activity of the gene and expression of PSA is largely mediated through androgen responsive elements (AREs) located in the proximal promoter and 5′ upstream enhancer.[6] Rao *et al*.[12] initially identified a single SNP within the region, a G/A substitution at −158 bp in the ARE I. It was subsequently reported that possession of the A allele was associated with increased serum PSA.[13,14] However, the associations between the ARE I polymorphism with either PSA levels have not been consistent.[15–17] In a study of 518 men, Xu *et al*. reported no association between polymorphisms in ARE I and PSA levels. The authors suggest that possession of the G allele was associated with increased serum PSA among white men participating in their study, but also note that the highest serum PSA levels among black subjects were observed with the AA genotype. Furthermore, Rao *et al*. found no difference in serum PSA associated with the ARE I genotype ($P=0.79$). Both studies also examined the relationship between the ARE I genotype and polymorphisms in the androgen receptor gene and with PSA and observed no significant interaction in contrast to other reports.[18]

The A allele at nucleotide position −158 has also been associated with prostate cancer risk in some studies,[14,19] but not in all.[17] It is worth noting that some studies report that overall risk appears higher among men with the AA genotype, and there are other publications showing that the GG genotype is associated with development of advanced cancer. Xue *et al*.[18] observed a higher frequency of the GG PSA genotype among men with advanced prostate cancer (evidence of extraprostatic extension, invasion into surrounding tissue, nodal involvement, and/or metastatic disease) (OR = 2.90; 95% CI = 1.24–6.78). A study conducted in China recently reported the GG PSA genotype was associated both with prostate cancer risk overall (OR = 2.27; $P = 0.008$) and also with greater tumor volume and higher pathologic stage.[20] The GG genotype has also been observed more frequently among patients with a Gleason grade of 7 and higher.[19] The inconsistency in reported observations with respect to the −158 SNP in ARE I of the PSA gene are due in part to differences in the racial and ethnic composition of eligible study subjects. For example, the GG genotype at the −158 SNP is seen in 24% of African-Americans men,[13] 29% of Non-Hispanic white men,[13] 61% of Japanese men,[17] and 62% of Chinese men.[20] This fact, coupled with relatively sample size of the study populations, may contribute to the differing results between the ARE I SNP and both serum PSA level as well as prostate cancer diagnosis. Importantly, the −158 SNP has been shown using *in vitro* assays to have no effect on PSA gene promoter activity.[12]

Using samples from a community-based study of African-American men from the FMHS, we elected to focus on the three PSA promoter SNPs (–4643 A/G, –5412 T/C, and –5429 T/G) upstream from the −158 A/G SNP discussed above. Using transfection assays into the human prostate cancer cell line LNCaP, luciferase reporter constructs containing either the G substitution at −4643 (along with a T at both position −5429 and −5412) or a combination of the C substitution at position −5412 and the G substitution at position −5429 (along with an A at position −4642) each showed more potent promoter activity using transient compared to constructs with the alternative nucleotide at the respective polymorphic site.[7] In addition to the *in vitro* evidence that these variants were functionally important, all three SNPs were shown by Cramer *et al*. to be in tight linkage disequilibrium with the −158 ARE I polymorphisms. However, in our study, we were unable to detect evidence that these genetic polymorphisms were associated with either total or free serum PSA levels. The observed frequency of the functional alleles of interest in our study population were approximately half (∼10 vs ∼20%) of that reported among white men in the Cramer study.

There are distinct differences in the design of each of these studies that might account for some of the inconsistency between findings. The primary aim of the investigation conducted by Cramer and co-workers was to examine the relationship between occupational asbestos exposure and lung disease among asbestos-exposed workers. Serum PSA was drawn initially to examine the relationship between asbestos exposure and prostate cancer risk and it is unknown if blood was drawn at the same time of the day for each participant. Evidence suggests that there is significant diurnal variability in PSA levels among men with and without prostate disease and the variation is unpredictable.[21,22] The FMHS protocol required blood draws for all participants between 0900 and 1100 hours to reduce the impact of temporal variability of results. Furthermore, because the participants of Cramer *et al*. were selected for inclusion into the study based upon their occupation, it is possible that some exposure unique to their occupation, whether asbestos or another substance, might alter their PSA levels. However, it is also possible that the findings represent real racial/ethnic differences in the prevalence and importance of these SNPs in influencing PSA and the utility of these SNPs in predicting PSA and/or prostate cancer risk in Caucasian populations may be greater than in African Americans. Further study is required before any firm conclusions can be drawn with respect to these SNPs in any population.

There are limitations in our study which must be discussed in the context of our findings. First, we have a relatively small sample size which places some limitations on our statistical power to detect small effects due to genotype. Specifically, given our observed genotype frequencies and standard errors in log10 PSA measurements by group, we would have less than 80% power to detect a significant difference (using $\alpha = 0.05$) in PSA level means between genotypes that is smaller than approximately 15%. Second, we did not genotype all SNPs within the PSA gene, therefore other SNPs within the gene may be associated with PSA serum levels in African Americans. We chose to focus on three functionally important SNPs in the promoter region in the PSA gene that have been reported previously as being associated with serum PSA levels. Third, among the control subjects, less than 60% of men who completed the initial epidemiologic survey participated in the blood draw and clinical examination. An analysis of the potential selection bias as a result of response rate in this population revealed that although men who participated in all phases of the study were younger, more likely to have a family history of prostate cancer and were likely to report the presence of urologic symptoms than nonparticipants, greater participation in the clinical phase of the study did not bias the estimated age-specific reference ranges for total PSA concentrations.[23] Based on these analyses, we have no reason to believe that participants and nonparticipants differ systematically with respect to their genetic background.

In conclusion, our analysis of 475 African-American men with and without prostate cancer does not suggest that SNPs in the 5′ upstream promoter region of the PSA gene influence free or total serum PSA level. Moreover, we did not demonstrate an association between these polymorphisms and prostate cancer risk. The observed prevalence of the alleles of interest in this community-based study were substantially lower than the prevalence reported in a study of Caucasian men published by Cramer *et al*. Given the strong linkage disequilibrium between these SNPs as observed in both studies, it will be difficult to establish the independent contribution of each polymorphism. Clearly, additional studies are needed to determine the importance of these SNPs as predictors of serum PSA levels and/or whether these polymorphisms are associated with prostate cancer risk. While enhanced methods for detecting prostate cancer in asymptomatic men are clearly required, it is not clear

that PSA promoter polymorphisms will be able to provide much improvement in the current standard of care using serum measurements of PSA.

Acknowledgements

We thank Julie McLaughlin for assistance with data analysis and Lisa Robbins for expert secretarial support. We also gratefully acknowledge support from the Department of Urology, University of Michigan Medical School; the University of Michigan Comprehensive Cancer Center Affymetrix and cDNA Microarray Core Facility; a Healthy Disparity Research Prostate Scholar Award from the Department of Defense, and the University of Michigan Prostate Cancer S.P.O.R.E. (CA069568).

References

1 Jemal A, Murray T, Ward E, Samuels A, Tiwari RC, Ghafoor A *et al*. Cancer Statistics 2005. *Ca Cancer J Clin* 2005; **55**: 10–30.
2 Bostwick DG, Burke HB, Djakiew D, Euling S, Ho SM, Landolph J *et al*. Human prostate cancer risk factors. *Cancer* 2004; **101**: 2371–2490.
3 Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, Redwine E. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. *N Engl J Med* 1987; **317**: 909–916.
4 Grossfeld GD, Carroll PR. Prostate cancer early detection: a clinical perspective. *Epidemiol Rev* 2001; **23**: 173–180.
5 Meitz JC, Edwards SM, Easton DF, Murkin A, Ardern-Jones A, Jackson RA *et al*. HPC2/ELAC2 polymorphisms and prostate cancer risk: analysis by age of onset of disease. *Br J Cancer* 2002; **87**: 905–908.
6 Kim J, Coetzee GA. Prostate specific antigen gene regulation by androgen receptor. *J Cell Biochem* 2004; **93**: 233–241.
7 Cramer SD, Chang BL, Rao A, Hawkins GA, Zheng SL, Wade WN *et al*. Association between genetic polymorphisms in the prostate-specific antigen gene promoter and serum prostate-specific antigen levels. *J Natl Cancer Inst* 2003; **95**: 1044–1053.
8 Lewontin RC. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* 1964; **49**: 49–67.
9 Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002; **70**: 425–434.
10 Cooney KA, Strawderman MS, Wojno KJ, Doerr KM, Taylor A, Alcser KH *et al*. Age-specific distribution of serum prostate-specific antigen in a community-based study of African-American men. *Urology* 2001; **57**: 91–96.
11 Diamandis EP, Yousef GM. Human tissue kallikrein gene family: a rich source of novel disease biomarkers. *Expert Rev Mol Diagn* 2001; **1**: 182–190.
12 Rao A, Chang BL, Hawkins G, Hu JJ, Rosser CJ, Hall MC *et al*. Analysis of G/A polymorphism in the androgen response element I of the PSA gene and its interactions with the androgen receptor polymorphisms. *Urology* 2003; **61**: 864–869.
13 Xue WM, Coetzee GA, Ross RK, Irvine R, Kolonel L, Henderson BE *et al*. Genetic determinants of serum prostate-specific antigen levels in healthy men from a multiethnic cohort. *Cancer Epidemiol Biomark Prev* 2001; **10**: 575–579.
14 Medeiros R, Morais A, Vasconcelos A, Costa S, Pinto D, Oliveira J *et al*. Linkage between polymorphisms in the prostate specific antigen ARE1 gene region, prostate cancer risk, and circulating tumor cells. *Prostate* 2002; **53**: 88–94.
15 Xu J, Meyers DA, Sterling DA, Zheng SL, Catalona WJ, Cramer SD *et al*. Association studies of serum prostate-specific antigen levels and the genetic polymorphisms at the androgen receptor and prostate-specific antigen genes. *Cancer Epidemiol Biomark Prev* 2002; **11**: 664–669.
16 Nam RK, Zhang WW, Trachtenberg J, Jewett MA, Emami M, Vesprini D *et al*. Comprehensive assessment of candidate genes and serological markers for the detection of prostate cancer. *Cancer Epidemiol Biomark Prev* 2003; **12**: 1429–1437.
17 Wang LZ, Sato K, Tsuchiya N, Yu JG, Ohyama C, Satoh S *et al*. Polymorphisms in prostate-specific antigen (PSA) gene, risk of prostate cancer, and serum PSA levels in Japanese population. *Cancer Lett* 2003; **202**: 53–59.
18 Xue W, Irvine RA, Yu MC, Ross RK, Coetzee GA, Ingles SA. Susceptibility to prostate cancer: interaction between genotypes at the androgen receptor and prostate-specific antigen loci. *Cancer Res* 2000; **60**: 839–841.
19 Gsur A, Preyer M, Haidinger G, Zidek T, Madersbacher S, Schatzl G *et al*. Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. *Carcinogenesis* 2002; **23**: 1647–1651.
20 Chiang CH, Chen KK, Chang LS, Hong CJ. The impact of polymorphism on prostate specific antigen gene on the risk, tumor volume and pathological stage of prostate cancer. *J Urol* 2004; **171**: 1529–1532.
21 Tekin A, Atsu N, Ozen H. Daily variability of serum prostate-specific antigen in men over 50 years of age. *Int Urol Nephrol* 2001; **33**: 641–644.
22 Dejter Jr SW, Martin JS, McPherson RA, Lynch JH. Daily variability in human serum prostate-specific antigen and prostatic acid phosphatase: a comparative evaluation. *Urology* 1988; **32**: 288–292.
23 Heeringa SG, Alcser KH, Doerr K, Strawderman M, Cooney K, Medbery B *et al*. Potential selection bias in a community-based study of PSA levels in African-American men. *J Clin Epidemiol* 2001; **54**: 142–148.

nature genetics

A common variant associated with prostate cancer in European and African populations

Laufey T Amundadottir[1,12], Patrick Sulem[1,12], Julius Gudmundsson[1,12], Agnar Helgason[1], Adam Baker[1], Bjarni A Agnarsson[2], Asgeir Sigurdsson[1], Kristrun R Benediktsdottir[2], Jean-Baptiste Cazier[1], Jesus Sainz[1], Margret Jakobsdottir[1], Jelena Kostic[1], Droplaug N Magnusdottir[1], Shyamali Ghosh[1], Kari Agnarsson[1], Birgitta Birgisdottir[1], Louise Le Roux[1], Adalheidur Olafsdottir[1], Thorarinn Blondal[1], Margret Andresdottir[1], Olafia Svandis Gretarsdottir[1], Jon T Bergthorsson[1], Daniel Gudbjartsson[1], Arnaldur Gylfason[1], Gudmar Thorleifsson[1], Andrei Manolescu[1], Kristleifur Kristjansson[1], Gudmundur Geirsson[3], Helgi Isaksson[2], Julie Douglas[4], Jan-Erik Johansson[5], Katarina Bälter[6], Fredrik Wiklund[6], James E Montie[7], Xiaoying Yu[8], Brian K Suarez[9], Carole Ober[10], Kathleen A Cooney[7,11], Henrik Gronberg[6], William J Catalona[8], Gudmundur V Einarsson[3], Rosa B Barkardottir[2], Jeffrey R Gulcher[1], Augustine Kong[1], Unnur Thorsteinsdottir[1] & Kari Stefansson[1]

With the increasing incidence of prostate cancer, identifying common genetic variants that confer risk of the disease is important. Here we report such a variant on chromosome 8q24, a region initially identified through a study of Icelandic families. Allele –8 of the microsatellite DG8S737 was associated with prostate cancer in three case-control series of European ancestry from Iceland, Sweden and the US. The estimated odds ratio (OR) of the allele is 1.62 ($P = 2.7 \times 10^{-11}$). About 19% of affected men and 13% of the general population carry at least one copy, yielding a population attributable risk (PAR) of ~8%. The association was also replicated in an African American case-control group with a similar OR, in which 41% of affected individuals and 30% of the population are carriers. This leads to a greater estimated PAR (16%) that may contribute to higher incidence of prostate cancer in African American men than in men of European ancestry.

The incidence of prostate cancer has increased markedly over the last decades and is now the most prevalent noncutaneous cancer in males in developed regions of the world[1]. The only firmly established risk factors for prostate cancer are age, family history of prostate cancer and ethnicity[2,3]. African Americans have among the highest incidence of prostate cancer and mortality rate attributable to this disease. African Americans are 1.6 times more likely to develop prostate cancer, and 2.4 times more likely to die from this disease, than European Americans[4]. Genetic factors probably contribute to such differences, along with a combination of environmental factors. However, in spite of compelling evidence for the role of genetics in prostate cancer[5], the search for common germline risk variants has been largely unrewarding[6].

In an attempt to identify genetic variants underlying risk of prostate cancer, we conducted a genome-wide linkage scan using 1,068 microsatellite markers typed for 871 Icelandic men with prostate cancer that grouped into 323 extended families (see Methods and the **Supplementary Note** online). This scan produced a suggestive linkage signal on chromosome 8q24 with a maximum lod score of 2.11 (D8S529 at 148.25 cM) (**Fig. 1a**).

To refine the source of this linkage signal, we genotyped an additional 358 microsatellite and indel markers spanning 18.6 cM on chromosome 8 from 125–135 Mb (NCBI Build 34) in 869 unrelated men with prostate cancer and 596 population controls (case-control group I) (**Fig. 1a**,**b**). We tested a total of 1,624 alleles. The strongest association to prostate cancer was observed for allele –8 of the microsatellite DG8S737, with an OR of 1.79 ($P = 3.0 \times 10^{-6}$) (**Fig. 1b** and **Table 1**). As population controls were used, the OR can also be considered an estimate of relative risk (RR) per copy carried based on the multiplicative model of risk[7,8]. This association was replicated in a second Icelandic case-con-

[1]deCODE genetics, Sturlugata 8, 101 Reykjavik, Iceland. [2]Department of Pathology, Landspitali-University Hospital of Iceland, 101 Reykjavik, Iceland. [3]Department of Urology, Landspitali-University Hospital of Iceland, 101 Reykjavik, Iceland. [4]Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA. [5]Department of Urology and Clinical Medicine, Örebro University Hospital, Sweden. [6]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, SE-17177 Stockholm, Sweden. [7]Department of Urology, University of Michigan, Ann Arbor, Michigan 48109, USA. [8]Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60611, USA. [9]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri 63110, USA. [10]Department of Human Genetics, University of Chicago, Chicago, Illinois 60637, USA. [11]Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan 48109, USA. [12]These authors contributed equally to this work. Correspondence should be addressed to K.S. (kstefans@decode.is) or U.T. (unnur.thorsteinsdottir@decode.is).

Received 9 March; accepted 25 April; published online 7 May 2006; doi:10.1038/ng1808

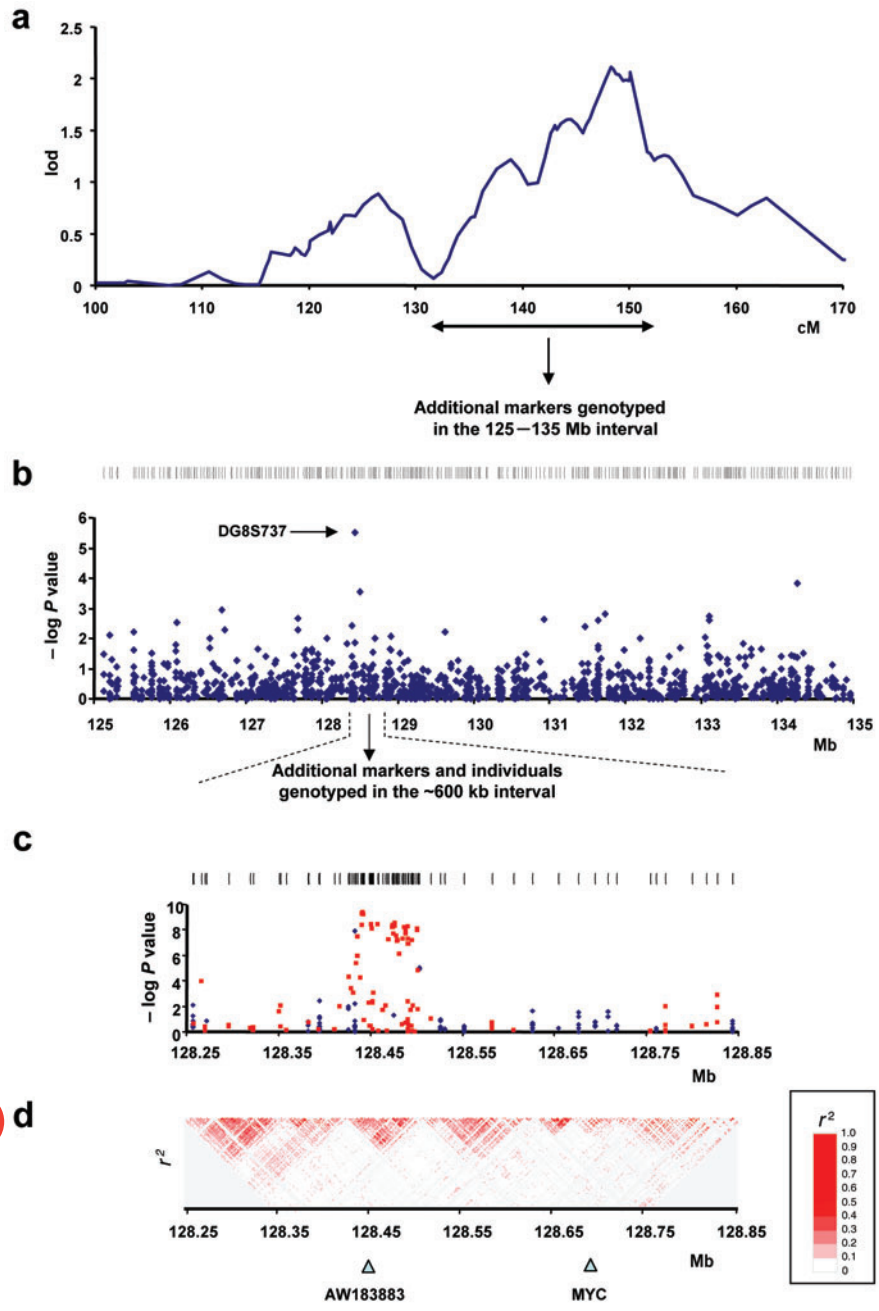

Figure 1 A schematic view of the linkage and association results, marker density and LD structure in a region on chromosome 8q24.21. (**a**) Linkage scan results for chromosome 8q performed with 871 Icelandic men with prostate cancer in 323 extended families. (**b**) Single-marker association results for unrelated prostate cancer cases (case-control group I, $n = 869$), using 358 microsatellites and insertions-deletions (indels) (blue diamonds), distributed over a 10-Mb region. (**c**) Single-marker association results for all prostate cancer cases ($N = 1291$). Filled red squares denote P values for the 63 SNPs and 12 microsatellites added to this region; blue diamonds denote the values for the other markers already typed in this region from **b**. (**d**) Pairwise LD from the CEU HapMap population for the 600 kb region from **c**. Filled gray triangles at the bottom indicate the location of the c-*MYC* gene and the AW183883 EST discussed in the main text. A scale for r^2 is provided on the right. Black vertical lines represent the density of microsatellites (in **b**) or microsatellites and SNPs (in **c**) used in the association analysis.

trol group of 422 men with prostate cancer and 401 population-based controls (case-control group II), in which the OR of allele –8 was 1.72 ($P = 0.0018$). In the combined Icelandic case-control groups I and II (1,291 affected men and 997 controls), the DG8S737 –8 allele had a frequency of 13.1% in affected men and 7.8% in controls ($OR = 1.77$, $P = 2.3 \times 10^{-8}$) (**Table 1**), corresponding to a PAR of 11%. DG8S737 (128.433096 Mb; **Supplementary Table 1** online) is a dinucleotide AC repeat located within a linkage disequilibrium (LD) block that spans 92 kb on chromosome 8q24.21 (128.414 to 128.506 Mb, NCBI Build 34) in the Utah CEPH (CEU) HapMap samples (all references to the HapMap data in the text are to release 19).

To further investigate the nature and extent of the association signal, we genotyped 63 SNPs and 12 additional microsatellites in a 600-kb region surrounding DG8S737 (**Fig. 1c,d** and **Supplementary Tables 1** and **2** online). In total, 53 SNPs and six microsatellites were located within the LD block that contains DG8S737, capturing most of its haplotype diversity, according to the CEU HapMap data. Of these SNPs, 37 were significantly associated with prostate cancer ($P < 0.001$), with allele A of SNP rs1447295 showing the strongest association ($OR = 1.72$, $P = 1.7 \times 10^{-9}$) (**Table 1**). Sixteen of the SNPs belong to the same LD equivalence class ($r^2 = 1$) as rs1447295 in the CEU HapMap sample and therefore showed comparable association. In the Icelandic samples, allele –8 of DG8S737 and allele A of rs1447295 were substantially correlated ($r^2 \approx 0.5$) (**Supplementary Table 3** online). This correlation was lower in the CEU HapMap sample ($r^2 \approx 0.3$), but no other SNP in the HapMap data had a higher correlation (**Supplementary Table 3**). In other words, the SNPs that were most associated with allele –8 of DG8S737 are also most associated with prostate cancer.

In an effort to identify new risk variants, we sequenced pools of DNA samples consisting of 117 Icelandic men with prostate cancer and 109 controls across the 92-kb LD block region (92% coverage) (**Supplementary Methods** online). Few new SNPs were identified, and none showed stronger association to prostate cancer in the combined Icelandic case-control group than allele –8 of DG8S737 or allele A of rs1447295.

We next attempted to replicate this association in 1,435 unrelated men with prostate cancer and 779 population-based controls from Sweden, and in 458 European American men with prostate cancer and 247 controls from Chicago. The frequency of the DG8S737 –8 allele was significantly greater in affected men than in controls for both the Swedish (OR 1.38, $P = 0.0043$) and European American (OR 2.10, $P = 0.0029$) case-control groups. We obtained a similar outcome for the rs1447295 A allele (**Table 1**), indicating that the variants initially identified in Iceland

Table 1 Association of alleles at chromosome 8q24 to prostate cancer in Iceland, Sweden and the US

Study population (*N* cases/*N* controls)	Marker	Allele	Allelic frequency		OR	*P* value
			Cases	Controls		
Iceland						
Group I[a]						
(869/596)	DG8S737	–8	0.134	0.080	1.79	3.0×10^{-6}
Group II[b]						
(422/401)	DG8S737	–8	0.124	0.076	1.72	1.8×10^{-3}
Combined groups I and II[b]						
(1,291/997)	DG8S737	–8	0.131	0.078	1.77	2.3×10^{-8}
	rs1447295	A	0.169	0.106	1.72	1.7×10^{-9}
Sweden[c]						
(1,435/779)	DG8S737	–8	0.101	0.079	1.38	4.3×10^{-3}
	rs1447295	A	0.164	0.133	1.29	4.5×10^{-3}
European Americans, Chicago						
(458/247)	DG8S737	–8	0.082	0.041	2.10	2.9×10^{-3}
	rs1447295	A	0.127	0.081	1.66	6.7×10^{-3}
African Americans, Michigan[b,d]						
(246/352)	DG8S737	–8	0.234	0.161	1.60	2.2×10^{-3}
	rs1447295	A	0.344	0.313	1.15	0.29

Shown are alleles for the markers DG8S737 and rs1447295 at 8q24.21 and the corresponding numbers of cases and controls (*N*), allelic frequencies of variants in affected and control individuals, odds ratio (OR) and two-sided *P* values.

[a]Individuals are unrelated at three meioses. [b]The association analysis was adjusted for the relatedness of some of the individuals. [c]The *P* values and the OR values were adjusted for a covariate that represented the two different centers (see Methods). [d]The results shown were adjusted for European and African ancestry between affected individuals and controls and were practically identical to unadjusted results.

are likely to be associated with increased risk of prostate cancer in most populations of European ancestry.

The rs1447295 A allele is more common than the DG8S737 –8 allele, and most chromosomes that carry the –8 allele also carry the A allele. Investigating the risk jointly (**Supplementary Table 4** online and **Supplementary Methods**), we found that the chromosomes that carry both the –8 allele and the A allele have the highest risk. The chromosomes that carry the A allele but not the –8 allele have lower risk than the former, but they still have significantly higher risk than chromosomes that carry neither allele (OR = 1.25, $P = 0.015$). These results indicate that neither the DG8S737 –8 nor the rs1447295 A alleles by themselves can fully explain the risk profile. Hence, either there are multiple functional variants in the region, or these alleles are both in strong, but imperfect, LD with a presently unknown risk variant. Accordingly, both markers should be genotyped in replication studies in populations of European ancestry.

We undertook a third replication study in 246 African American men with prostate cancer and 352 controls to determine whether the variants identified above are also associated with prostate cancer in a group with high incidence of the disease. If so, then the greater genetic diversity in African Americans could provide greater resolution to pinpoint the unknown risk variant. This assumption was supported by an analysis of the 92-kb LD block in the Nigerian Yoruba (YRI) HapMap sample, which uncovered both greater genetic diversity and weaker LD in the YRI sample than in populations of European ancestry. Specifically, although 19 SNPs, including rs1447295, are in the same equivalence class ($r^2 = 1$) in the CEU HapMap data, they belong to 13 different equivalence classes in the HapMap YRI sample (**Supplementary Table 3**). Consequently, in addition to DG8S737, we genotyped 17 of the 19 equivalent SNPs (including rs1447295) in the African American case-control group. Of the two omitted, one was perfectly correlated with two other SNPs we genotyped, and the other was nonpolymorphic, in the YRI samples.

Allele frequency differences between the YRI HapMap sample and the groups of European ancestry raised the possibility that spurious association results could arise from differences in the distribution of European ancestry among the African American men in the affected and control groups. To control for ancestry, we genotyped 30 microsatellites that are randomly distributed in the genome and informative for distinguishing between African and European ancestry. An analysis of these data with Structure[9–11] did not show any significant differences in European ancestry between affected individuals and controls.

The frequency of allele –8 of DG8S737 was 23.4% in African American men with prostate cancer and 16.1% in controls, with an OR of 1.60 ($P = 0.0022$). Of the 17 SNPs that we genotyped, rs1447295 gave the lowest,

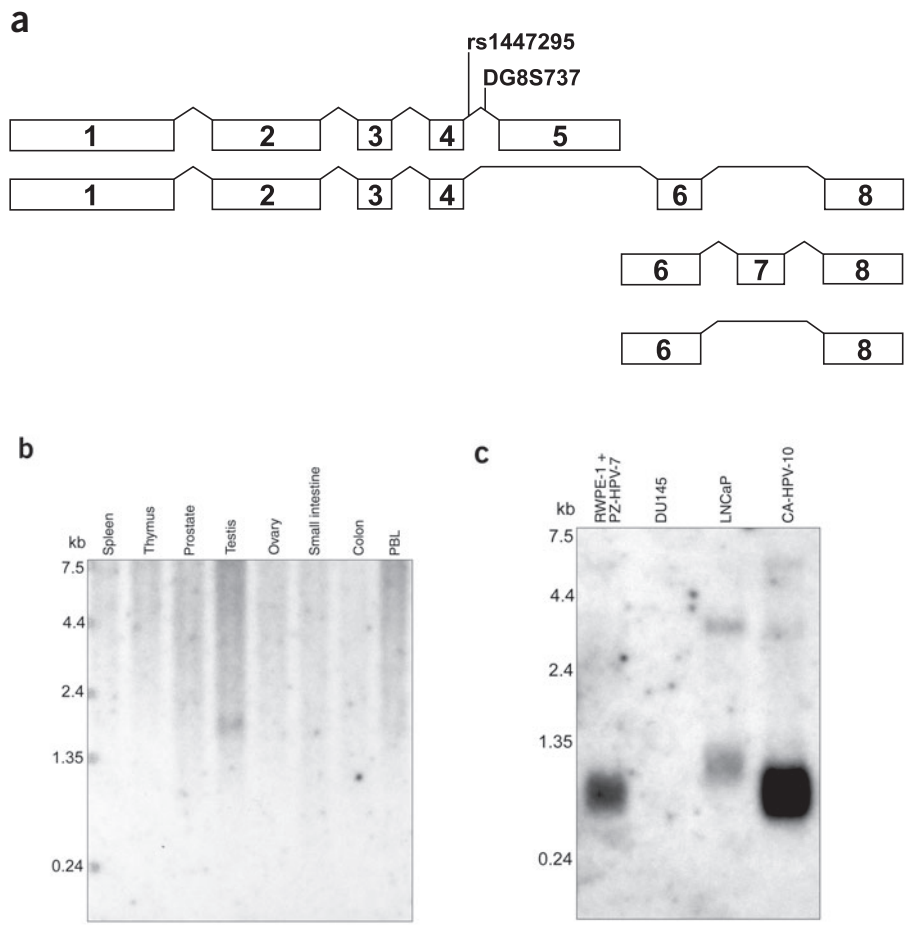

Figure 2 RNA expression of identified splice variants derived from the EST AW 183883. (**a**) Schematic representation of the four AW splice variants identified. Exons are shown as rectangles and introns as lines. The transcripts extend from 128.258–128.451 Mb on chromosome 8q24. The length of exons is as follows: exon 1: 503 bp; exon 2: 343 bp; exon 3: 103 bp; exon 4: 88 bp; exon 5: 371 bp; exon 6: 135 bp; exon 6 (long): 546 bp; exon 7: 140 bp and exon 8: 246 bp. Shown are locations of the microsatellite marker DG8S737 and the SNP rs1447295 relative to the splice variants. Figure is not drawn to scale. (**b**) RNA blot analysis of polyA RNA isolated from various human tissues. The membrane was hybridized with the 351-bp full-length AW183883 EST cDNA (exons 1 and 2). PBL, peripheral blood leukocyte. Note the 1.5-kb band in testis. (**c**) RNA blot analysis of human polyA RNA from normal and malignant prostate cell lines. The membrane was hybridized with a 657-bp PCR-amplified probe corresponding to exons 6–8 (see **Supplementary Methods** for sequence). Note the ~600 bp band in the RWPE-1/PZ-HPV-7 and CA-HPV-10 lines and the ~900-bp band in the LNCaP cell line. The RWPE-1 and PZ-HPV-7 cell lines are derived from normal prostate and the others from prostate cancer. The membrane in **b** was also hybridized with a probe specific for exon 6–8 but gave no signal. Similarly, the AW183883 EST cDNA (exons 1 and 2) probe did not give a signal when hybridized to the membrane in **c**.

albeit not significant, *P* value ($P = 0.29$) (**Table 1**). This indicates that DG8S737 –8, rather than the SNPs, is either itself a functional variant or is very tightly associated with a presently unknown risk variant both in populations of European and of African ancestry. Although the latter possibility might seem unlikely given the considerable range of allele sizes exhibited by DG8S737 (**Supplementary Table 5** online), a phylogenetic analysis indicated that this marker has only a moderate mutation rate, with strong correlation between allele size and SNP haplotype background in the HapMap samples (**Supplementary Fig. 1** online).

Assessment of the HapMap YRI data showed that the three SNPs that are most strongly correlated with the –8 allele of DG8S737 ($r^2 = 0.32$ to 0.34) were among the 17 SNPs genotyped in the African American samples (**Supplementary Table 3**). Thus, it is highly unlikely that any SNP within the LD block in the HapMap data will show a strong association to prostate cancer in African Americans. Even though the OR is similar in the case-control groups of African and European ancestry, the estimated PAR in African Americans is considerably higher (16% versus 5%–11%) because of the greater frequency of DG8S737 –8. This allele is even more common in the YRI HapMap samples (22.5%), raising the possibility that the PAR captured by DG8S737 –8 may even be greater in African populations. We found that the multiplicative risk model fitted the data adequately for populations of both European and African ancestry. Genotype counts and model-free estimates of genotype-specific RR values can be found in **Supplementary Tables 6** and **7** online.

We next sought to determine whether the risk variants associate more strongly with aggressive forms of prostate cancer as reflected by high Gleason scores. In all four case-control groups, the frequency of DG8S737 –8 was significantly greater in men with prostate cancer, with combined Gleason scores of 7 to 10, than in controls (**Table 2**). Moreover, the frequency of allele –8 was greater in affected individuals with high (7–10) than with low (2–6) Gleason scores in all four case-control groups combined ($OR = 1.21$, $P = 0.02$) and the three European ancestry case-control groups combined ($OR = 1.18$, $P = 0.07$), but the difference is modest. An analysis of 510 Icelandic men diagnosed with benign prostatic hyperplasia (BPH), but not prostate cancer, showed no significant excess of either allele –8 of DG8S737 or allele A of rs1447295 (data not shown) indicating that these variants increase the risk of malignant prostate tumors only, and might have a stronger association with the more aggressive forms.

As only the microsatellite allele showed significant association in the African American case-control group, and it is contained in a smaller LD block in African Americans than in populations of European ancestry (**Supplementary Fig. 2** online), we propose that the region most likely to contain the functional variant can be narrowed down to positions 128.414–128.474 Mb (NCBI build 34). This region contains one spliced EST (AW183883) and three single-exon ESTs (BE144297, CV364590 and AF119310) in addition to a few predicted genes, but no known genes[12]. No microRNAs have been detected within the block[13]. Expression analysis in various cDNA libraries confirmed only the expression of the AW183883 EST (**Supplementary Methods**). We identified four different splice variants of AW183883 by 5′and 3′ RACE that were verified by RT-PCR and RNA blot analysis (**Fig. 2a**). Using the AW183883 EST as a probe on an RNA blot, we detected a ~1.5-kb signal only in testis, consistent with the size of the two longer forms. The two shorter transcripts harboring exons 6–8 were detected only in normal (0.6-kb transcript) and malignant (0.6- and 0.9-kb transcripts) prostate cell lines, not in

Table 2 Association of alleles at chromosome 8q24 to high and low Gleason scores in Iceland, Sweden and the US

Study population (*N* cases/ *N* controls)	Marker	Allele	Allelic frequency		OR	*P* value
			Cases	Controls		
Iceland[b]						
Biopsy Gleason 7–10						
(289/997)	DG8S737	–8	0.146	0.078	2.00	4.0×10^{-6}
	rs1447295	A	0.179	0.106	1.84	7.3×10^{-6}
Biopsy Gleason 2–6						
(548/997)	DG8S737	–8	0.131	0.078	1.78	3.4×10^{-6}
	rs1447295	A	0.170	0.106	1.73	6.7×10^{-7}
Sweden[a]						
Gleason 7–10						
(625/779)	DG8S737	–8	0.107	0.079	1.42	1.0×10^{-2}
	rs1447295	A	0.167	0.133	1.29	2.0×10^{-2}
Gleason 2–6						
(678/779)	DG8S737	–8	0.094	0.079	1.31	4.5×10^{-2}
	rs1447295	A	0.158	0.133	1.25	3.4×10^{-2}
European Americans, Chicago						
Biopsy Gleason 7–10						
(149/247)	DG8S737	–8	0.108	0.041	2.83	4.4×10^{-4}
	rs1447295	A	0.151	0.081	2.03	2.7×10^{-3}
Biopsy Gleason 2–6						
(306/247)	DG8S737	–8	0.071	0.041	1.78	3.6×10^{-2}
	rs1447295	A	0.116	0.081	1.50	5.1×10^{-2}
African Americans, Michigan[b,c]						
Biopsy Gleason 7–10						
(112/352)	DG8S737	–8	0.273	0.161	1.96	3.3×10^{-4}
	rs1447295	A	0.352	0.313	1.19	0.28
Biopsy Gleason 2–6						
(121/352)	DG8S737	–8	0.211	0.161	1.40	8.2×10^{-2}
	rs1447295	A	0.341	0.313	1.14	0.43

Shown are alleles for the markers DG8S737 and rs1447295 at 8q24.21 and the corresponding numbers of cases and controls (N), frequencies of variants in affected and control individuals, odds ratio (OR) and two-sided *P* values. About 80% Swedish Gleason scores are from biopsy material and the rest from surgery.

[a]The *P* values and the ORs were adjusted for a covariate that represented the two different centers (see Methods). [b]The association analysis was adjusted for the relatedness of some of the individuals. [c]The results shown were adjusted for European and African ancestry between affected individuals and controls and were practically identical to unadjusted results.

the other tissues analyzed (**Fig. 2b**,**c**). The predicted ORFs for these transcripts did not show significant homology to known proteins.

We note that 8q24 is the most frequently gained chromosomal region in prostate tumors[14]. Gain in this region has been associated with aggressive tumors, hormone independence and poor prognosis[15]. However, DNA blot analysis of the 92-kb LD region using germline-derived ($N = 31$) and tumor-derived ($N = 15$) DNA showed no differences between carriers and noncarriers of the DG8S737 –8 allele (data not shown). Also notable is the proximity of DG8S737 to the well-known oncogene *c-MYC* (~270 kb telomeric). However, we did not observe any significant correlation between SNPs located in the *c-MYC* gene and either prostate cancer risk or the risk variants identified in this study (data not shown). Nevertheless, it is possible that the risk variant modifies *c-MYC* regulation by predisposing to genomic instability or altering long-range regulation of expression, although we have not uncovered any evidence supporting these possibilities.

In summary, we have demonstrated significant association of prostate cancer risk to the DG8S737 –8 and rs1447295 A alleles in three case-control groups of European ancestry (in which rs1447295 A is perfectly correlated with alleles from at least 18 other nearby SNPs). Combining results from these groups gave an estimated OR of 1.62 ($P = 2.7 \times 10^{-11}$) for DG8S737 –8 and an OR of 1.51 ($P = 1.0 \times 10^{-11}$) for rs1447295 A. Assuming population frequencies of 6.6% and 10.7% (averages from the three groups), and using OR as an estimate of RR, the corresponding PAR was approximately 8% and 10%, respectively, for these two markers. The association between prostate cancer and the DG8S737 –8 allele was also replicated in an African American case-control group with nearly identical OR, but association in this group was not demonstrated for the HapMap SNPs that showed the strongest correlation in the CEU and the YRI samples with the –8 allele of DG8S737.

We identified the risk variants described here through a positional cloning approach, starting with linkage analyses. However, the association signal could have been detected, in theory, by genome-wide association in populations of European ancestry using common SNPs through either rs1447295 or one of its LD equivalents. Based on the *P* value from the combined European ancestry case-control groups, the result would

remain highly significant even if it were necessary to adjust for the testing of hundreds of thousands of common SNPs. In contrast, if based solely on SNPs contained in release 19 of the HapMap project, our analyses suggest that a genome-wide association study would not capture this association signal in African American or African samples. Consequently, we postulate that either the –8 allele confers risk, or it is more closely correlated with the risk variant than any of the current HapMap SNPs. If the latter hypothesis is true, then the lesser LD in African Americans indicates that the unknown variant is located within a 60-kb region of strong LD containing DG8S737. Also noteworthy is the relatively high population frequency, in African Americans, of the –8 allele, which confers an estimated PAR of about 16% and could alone produce more than a 10% greater incidence of prostate cancer in African Americans than in European Americans. Thus, this genetic variant might account, in part, for the higher incidence of prostate cancer in African Americans.

METHODS

Icelandic study population. Men with prostate cancer were identified based on a nationwide list from the Icelandic Cancer Registry containing all 3,815 Icelandic men with prostate cancer diagnosed from 1 January 1955 to 31 December 2004. A total of 1,291 affected individuals were recruited for this study out of the 1,425 affected individuals who were alive from November 2000 to June 2005 (a participation rate of 90.6%). For the linkage analysis, the deCODE genetics genealogical database[16] was used to group 871 affected individuals into 323 informative pedigrees. All 1,291 affected individuals were used in the case-control association analysis, but they were divided into two groups. Case-control group I (869 cases and 596 controls) consisted of unrelated individuals (at three meiosis). Some of the cases in case-control group II (422 cases and 401 controls) were related within three meiosis (28 relationships), and the same was true for some cases in the combined group of 1,291 cases and 997 controls (140 relationships). In the association analysis, we adjusted for the relatedness in these groups as described below. The 997 controls were randomly selected from the Icelandic genealogical database (see **Supplementary Note** for a more detailed description of the Icelandic study population).

The study was approved by the Data Protection Commission of Iceland and the National Bioethics Committee of Iceland. Written informed consent was obtained from all affected individuals, relatives and controls. Personal identifiers associated with medical information and blood samples were encrypted with a third-party encryption system as previously described[17].

Swedish and US study populations. CAPS1 ('Cancer of the Prostate in Sweden 1') is a population-based case-control study with enrollment from January 2001 through September 2002 (ref. 18). Genotypes from 1,435 Swedish men with prostate cancer and 779 controls were used in this study. Informed consent was obtained from all subjects, and the study was approved by the Ethics Committees at the Karolinska Institute and Umea University.

The European American study population consisted of 458 men with prostate cancer, recruited from the Pathology Core of Northwestern University's Prostate Cancer Specialized Program of Research Excellence (SPORE) from May 2002 to June 2005. The 247 European American controls were recruited as healthy control subjects for genetic studies at the University of Chicago and Northwestern University Medical School. Study protocols were approved by the Institutional Review Boards of Northwestern University and the University of Chicago. All subjects gave written informed consent.

The African American study population consisted of 352 male controls recruited through the Flint Men's Health Study (FMHS) and 246 men with prostate cancer from FMHS and the University of Michigan Prostate Cancer Genetics Project (PCGP)[19,20]. Affected individuals and controls were of self-reported African American ethnicity. Informed consent was obtained from all study participants, and protocols were approved by the Institutional Review Board at the University of Michigan Medical School (see **Supplementary Note** for more detailed descriptions of the Swedish and US study populations).

Linkage analysis. A genome-wide linkage scan was performed in Iceland with a framework set of 1,068 microsatellites. For multipoint linkage analysis, we used an affected-only allele-sharing method[21] as implemented in the program Allegro[22] and as previously described[23]. Our genetic map[24] was used. An additional 25 markers were typed in the region of suggestive linkage to increase the information content.

Association analysis. For single-marker association to prostate cancer, we used a likelihood ratio test to calculate a two-sided *P* value for each allele. Each allele was tested separately with the complementary alleles (when there was more than one complementary allele, they were collapsed into a single composite allele). We attempted to genotype all individuals reported in **Table 1** for DG8S737 and rs1447295. For both markers, yield was higher than 90% in every group (see **Supplementary Table 6** for details), and every individual had a genotype for at least one of the two markers. As the two markers were in strong LD, when the genotype of one marker was missing for an individual, the genotype of the other marker was used to provide partial information through a likelihood approach. This ensured that results for both markers presented in **Tables 1** and **2** were always based on the same individuals, allowing meaningful comparisons. A likelihood procedure described in a previous publication[8], implemented in the NEMO software, was used for haplotype analyses (see **Supplementary Methods** for details). We tested the association of an allele to prostate cancer using the signed (+ for excess in affected individuals; – for deficit) square root of a standard likelihood ratio statistic that, if the subjects were unrelated, would have asymptotically a standard normal distribution under the null hypothesis.

Allelic frequencies rather than carrier frequencies are presented for the markers. Allele-specific OR and RR were calculated assuming a multiplicative model[7]. Results from multiple case-control groups were combined using a Mantel-Haenszel model[25] in which the groups were allowed to have different population frequencies for alleles, haplotypes and genotypes but were assumed to have common relative risks. As described above, the case-control samples from Sweden came from two centers that have individually been frequency matched by geography and age. The results for the Swedish case-control study as a whole, presented in **Tables 1** and **2**, were also calculated using a Mantel-Haenszel model to adjust for a covariate that represents the different centers. For each of the four case-control groups, there was no significant deviation from Hardy-Weinberg equilibrium (HWE) in the controls.

Correction for relatedness. Some of the individuals with prostate cancer in the Icelandic case-control groups were related to each other (see above). The genotypes of closely related individuals are not independent, causing the standard deviation of the aforementioned association test statistic to be >1, which, if not corrected for, would lead to *P* values that are anticonservative. An adjustment for relatedness was performed using a previously described procedure[26]. We simulated 10,000 sets of genotypes for the marker DG8S737 through the genealogy of 708,683 Icelanders. With each simulated set, we recalculated the statistic by treating the simulated genotypes as real genotypes of the affected individuals and controls in the study. From the simulations, the true standard deviation of the statistic under the null hypothesis was 1.018 for allele –8, and this value was used to calculate the *P* values for the Icelandic total case-control group of 1,291 men with prostate cancer and 997 controls. Based on similar simulations, the adjustment factor for allele A of rs1447295 was found to be slightly lower, as expected, owing to the higher frequency of allele A compared to allele –8. We decided to use the higher adjustment factor of 1.018 throughout for simplicity. Hence, the results reported for allele A are slightly conservative. A similar procedure was used to adjust for the relatedness of some men with prostate cancer in the Michigan African American case-control group. Apart from the given relationships among some individuals in the PCGP African American Michigan study, further investigations with RELPAIR[27,28] using the ethnicity marker genotypes uncovered 14 additional cryptic relationships, 11 of them first-degree and three second-degree. Among the first-degree relations, seven are within groups (five case-case and two control-control) and four are between groups (case-control). The reported associations results were adjusted for both the known relations and the cryptic relations.

Evaluation of genetic ancestry. We used the program Structure[10] to estimate the genetic ancestry of individuals. Structure infers the allele frequencies of *K* ancestral populations on the basis of multilocus genotypes from a set of individuals and a user-specified value of *K* and assigns a proportion of ancestry from each

of the inferred K populations to each individual. The analysis of our data set was run with $K = 3$, with the aim of identifying the proportion of African and European ancestry in each individual. The statistical significance of the difference in mean European ancestry between African Americans with prostate cancer and controls was evaluated by reference to a null distribution derived from 10,000 randomized data sets.

To evaluate genetically estimated ancestry of the case-control groups from the US, we selected 30 unlinked microsatellite markers from about 2,000 microsatellites genotyped in a previously described[10] multiethnic cohort of 35 European Americans, 88 African Americans, 34 Chinese and 29 Mexican Americans. Of the 2,000 microsatellite markers, the selected set showed the most significant differences between European Americans, African Americans and Asians and also had good quality and yield.

Accession codes. GenBank: newly identified splice variants derived from the AW183883 EST: DQ515896 (exons 1-2-3-4-5), DQ515897 (exons 1-2-3-4-6-8), DQ515898 (exons 6-7-8) and DQ515899 (exons 6–8).

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank the men with cancer and their family members whose contribution made this work possible. We also thank the nurses at Noatun and personnel at the deCODE core facilities for their hard work and enthusiasm. In addition, we thank the following individuals at the Department of Pathology at Landspitali University-Hospital for their contributions: A. Arason, G. Johannesdottir, K. Olafsdottir and S. Kristjansdottir. We also thank the staff at the Icelandic Cancer Registry for providing patient lists. We thank Regional Cancer Registries in Umeå, Uppsala, Stockholm-Gotland and Lindköping and all urologists who recruited their patients to this study and provided clinical data to the national registry of prostate cancer. We also thank P. Stattin, J. Adolfsson and E. Varenhorst as regional coordinators in CAPS and the personnel at the Medical Biobank in Umeå. The CAPS study has been supported by grants from the Swedish Cancer Foundation and partially funded by a US National Cancer Institute grant (CA 1R01CA105055-01A1). We also thank participants and clinicians at Northwestern University, the University of Chicago and the University of Michigan for their contribution to the collection of samples and clinical data used in this study.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the *Nature Genetics* website for details).

Published online at http://www.nature.com/naturegenetics

1. Parkin, D.M., Bray, F., Ferlay, J. & Pisani, P. Global cancer statistics, 2002. *CA Cancer J. Clin.* **55**, 74–108 (2005).
2. Amundadottir, L.T. *et al.* Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. *PLoS Med* **1**, e65 (2004).
3. Crawford, E.D. Epidemiology of prostate cancer. *Urology* **62**, 3–12 (2003).
4. NIH Pub. No. 99–4649. Cancer incidence and survival among children and adolescents: United States SEER program 1975–1995 (US National Cancer Institute, SEER Program, Bethesda, Maryland, 1999).
5. Lichtenstein, P. *et al.* Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **343**, 78–85 (2000).
6. Schaid, D.J. The complex genetic epidemiology of prostate cancer. *Hum. Mol. Genet.* **13**, R103–R121 (2004).
7. Falk, C.T. & Rubinstein, P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann. Hum. Genet.* **51**, 227–233 (1987).
8. Gretarsdottir, S. *et al.* The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. *Nat. Genet.* **35**, 131–138 (2003).
9. Helgadottir, A. *et al.* A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nat. Genet.* **38**, 68–74 (2006).
10. Pritchard, J.K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
11. Pritchard, J.K., Stephens, M., Rosenberg, N.A. & Donnelly, P. Association mapping in structured populations. *Am. J. Hum. Genet.* **67**, 170–181 (2000).
12. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
13. Griffiths-Jones, S. The microRNA registry. *Nucleic Acids Res.* **32**, D109–D111 (2004).
14. Baudis, M. & Cleary, M.L. Progenetix.net: an online repository for molecular cytogenetic aberration data. *Bioinformatics* **17**, 1228–1229 (2001).
15. El Gedaily, A. *et al.* Discovery of new DNA amplification loci in prostate cancer by comparative genomic hybridization. *Prostate* **46**, 184–190 (2001).
16. Gulcher, J., Kong, A. & Stefansson, K. The genealogic approach to human genetics of disease. *Cancer J.* **7**, 61–68 (2001).
17. Gulcher, J.R., Kristjansson, K., Gudbjartsson, H. & Stefansson, K. Protection of privacy by third-party encryption in genetic research in Iceland. *Eur. J. Hum. Genet.* **8**, 739–742 (2000).
18. Lindmark, F. *et al.* H6D polymorphism in macrophage-inhibitory cytokine-1 gene associated with prostate cancer. *J. Natl. Cancer Inst.* **96**, 1248–1254 (2004).
19. Beebe-Dimmer, J.L. *et al.* Polymorphisms in the prostate-specific antigen gene promoter do not predict serum prostate-specific antigen levels in African-American men. *Prostate Cancer Prostatic Dis.* **9**, 50–55 (2006).
20. Douglas, J.A. *et al.* Identifying susceptibility genes for prostate cancer–a family-based association study of polymorphisms in CYP17, CYP19, CYP11A1, and LH-beta. *Cancer Epidemiol. Biomarkers Prev.* **14**, 2035–2039 (2005).
21. Kong, A. & Cox, N.J. Allele-sharing models: LOD scores and accurate linkage tests. *Am. J. Hum. Genet.* **61**, 1179–1188 (1997).
22. Gudbjartsson, D.F., Jonasson, K., Frigge, M.L. & Kong, A. Allegro, a new computer program for multipoint linkage analysis. *Nat. Genet.* **25**, 12–13 (2000).
23. Gretarsdottir, S. *et al.* Localization of a susceptibility gene for common forms of stroke to 5q12. *Am. J. Hum. Genet.* **70**, 593–603 (2002).
24. Kong, A. *et al.* A high-resolution recombination map of the human genome. *Nat. Genet.* **31**, 241–247 (2002).
25. Mantel, N. & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl. Cancer Inst.* **22**, 719–748 (1959).
26. Grant, S.F. *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323 (2006).
27. Boehnke, M. & Cox, N.J. Accurate inference of relationships in sib-pair linkage studies. *Am. J. Hum. Genet.* **61**, 423–429 (1997).
28. Epstein, M.P., Duren, W.L. & Boehnke, M. Improved inference of relationship for pairs of individuals. *Am. J. Hum. Genet.* **67**, 1219–1231 (2000).